US 20090044293A1

(54) **PLANTS WITH ALTERED ROOT ARCHITECTURE, INVOLVING THE RT1 GENE, RELATED CONSTRUCTS AND METHODS**

(76) Inventors: **Hajime Sakai**, Newark, DE (US); **Graziana Taramino**, Wilmington, DE (US)

Correspondence Address:
E I DU PONT DE NEMOURS AND COMPANY LEGAL PATENT RECORDS CENTER BARLEY MILL PLAZA 25/1122B, 4417 LANCASTER PIKE WILMINGTON, DE 19805 (US)

Publication Classification

(57)   **ABSTRACT**

Isolated polynucleotides and polypeptides and recombinant DNA constructs particularly useful for altering root structure of plants, compositions (such as plants or seeds) comprising these recombinant DNA constructs, and methods utilizing these recombinant DNA constructs. The recombinant DNA construct comprises a polynucleotide operably linked to a promoter functional in a plant, wherein said polynucleotide encodes a polypeptide useful for altering plant root architecture.
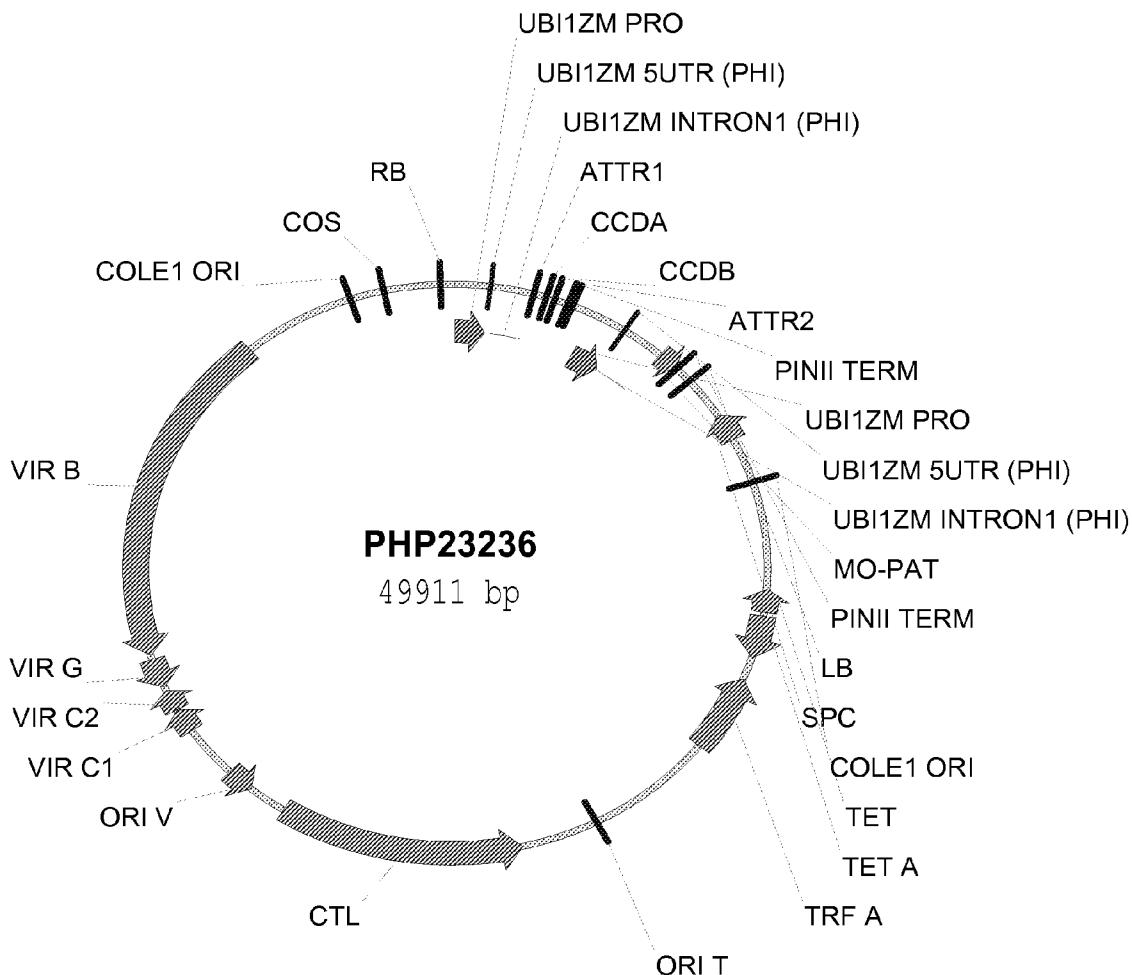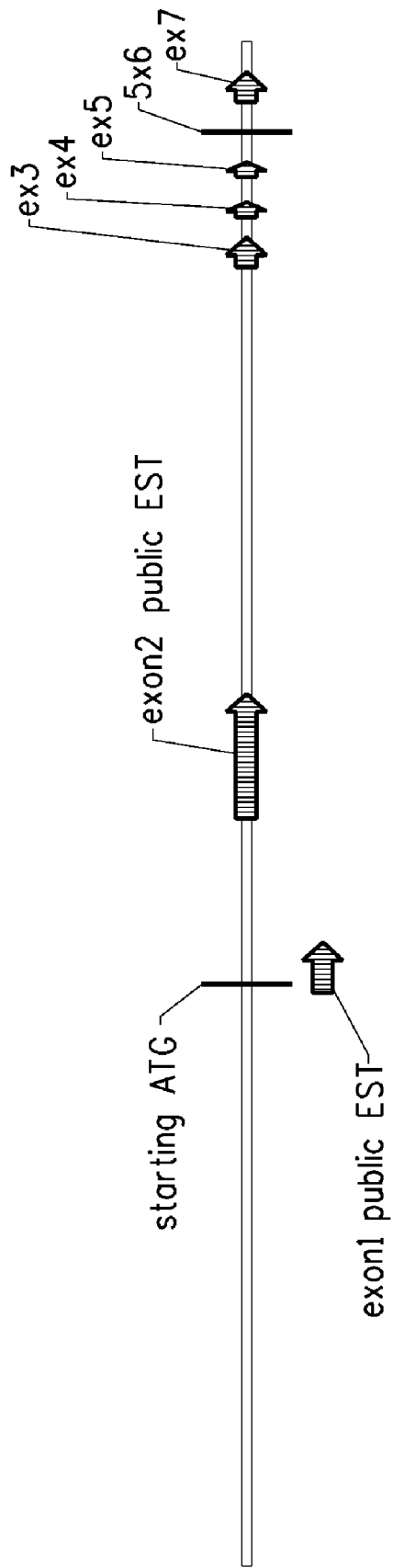
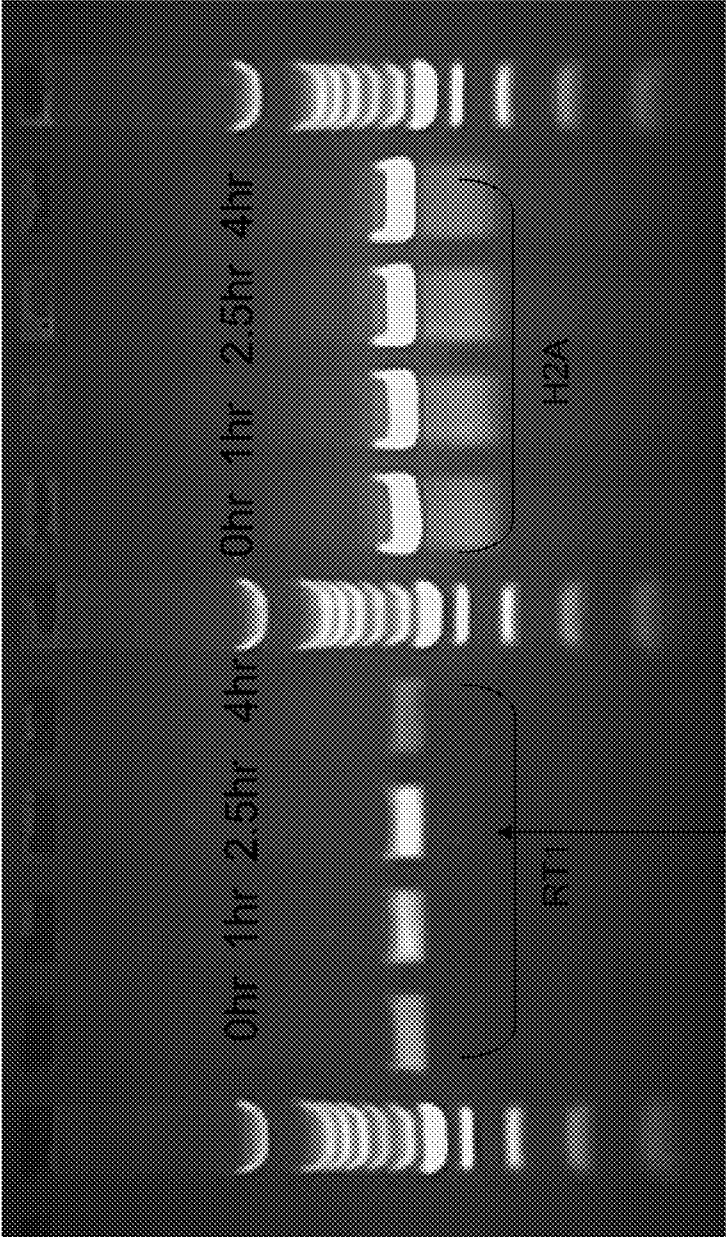New rt1 bac region around ethylen respo 7.8 Kb exon annotation
7800 bp

F I G. 1

FIG. 2

RT PCR after induction with Ethephon (1.5×10^{-4} M)

Induction after 2.5 Hours of treatment

FIG. 3A

```
                    **                   *        *  ** *                                          **
SEQ ID NO:13  MMAASSSSSFNLSDHLVQDDL-PWPASSSSSSQLPFAPSHPPHGAIIGGSQQWWQQPQML
SEQ ID NO:16  MMAASSSSS-NLSDHLAQDDLLPWPSSSSSS--LPFAPAAPHHSAIIGGNSQ---QPLAL
SEQ ID NO:19  MMASSSSSSLCDNHLLQDDLIPWPS------MPFAPAP---NTFGLNHQWSQPPMLS
SEQ ID NO:21  MDSA--------NLHQLQDQL-QLVGSSSSSSLDNNSDPSCYGA--SSAHQWSPGGISL
                                                                  *  **

SEQ ID NO:13  NCADRHPSDELEVLLSAQGSHHH--------HHHPSSSVIHPQLSS-LLMMQDLGFQWSSCG
SEQ ID NO:16  NCADRHHSDELEVLLSAQGSHHHQSHSHRHSHPASPVIPPQLSSSLLTMQDLGFQWSNCG
SEQ ID NO:19  SSTDQLSSYELESLQSVESQLAAAAPP---TLSPHLQAHQLSTVLMMQELGFQWSSCA
SEQ ID NO:21  NSVSLSHNYNNEMLNT-----RAHNNNNNNTSECMSLSSIHNHSLIQQQDFPLQWPHDQ

SEQ ID NO:13  S-FVDTSS-------SSPMPPPLVNGQQHDGHNKIKEWDRPLTAN---NSSRSSCASTAM---
SEQ ID NO:16  GGFLDTS--------PSTPPTNGQLQHDGHDDDKIEAEGTLIT---NSSRPSCAGTATIAA
SEQ ID NO:19  A------PA---DQHSIASSTNNNSNVMMNEEELRPRPDQSL---ISNPRSCSATTLLPP
SEQ ID NO:21  SSYQHHEGLLKIKEELSSSTISDHQEGISKFTDMLNSPVITNYLKINEHKDYTEKLLLKS

SEQ ID NO:13  ACRDVVLDGGGGGGGLPAMAAVAADLDD---GTVLPASVNVPRTTTTRQRPFPAAPPPLPG
SEQ ID NO:16  SCHDVVLHGGGGGG-TVLPTINISLTT---TTTHTQAQ---QQQRPFLAAPPPLPG
SEQ ID NO:19  P--HLHLDGA------VLPSINVS-------RLQKLAAGDEPL--
SEQ ID NO:21  MSSGFPINGDYGSSLPSSSSSSSPSSQSHRGNFSQIYPSVNISSLSESRKMSMDDMSNIS

SEQ ID NO:13  DAFEIL-ASSRLCKTLLLSQA-------SSSVLLHNG---TTMPLLRSE--HVPYG------
SEQ ID NO:16  DAFEILLASSRLCKTLLLSQAAAASSSSSSVLLHDDGTISTVPSLRSESEHVAYGYGGGPP
SEQ ID NO:19  -------QICCK---RQAAAAA-------VVGHSSIRD--EHVPCPYAGPP-
SEQ ID NO:21  RPFDINM--QVFDGRLF--------EGNVLVPPFNAQEISSLGMSRGSLPS

SEQ ID NO:13  PPAPA---HPQG--PSIGNYKQCMQMGAS-CGG-------RAQHEAACQSMAAAAARP
SEQ ID NO:16  PPAPAAAHRPRGRRPSVDNYKQ--QMGAAACAAGAGRQWSAEHGGRAAACEERAAAAPAR
SEQ ID NO:19  -----AHLIQGPSNT------LQMKRNTNAAAQGRG-----GRHGSSTEHRSSTA--
SEQ ID NO:21  FGLPFHHHLQQTLPHLSSSPTHQMEMFSNEPQTSEGK------RHNFLMATKAGENAS-
```

FIG. 3B

```
              ****  *              *  * ***********************************
SEQ ID NO:13  ALPVVPSSKKPRVEQRSS------------TIVPSFKVRKEKLGDRIAALQQLVSPFGKTDTA
SEQ ID NO:16  R-PVVVPWKKPRLEQRSGSGSGSGSSTTILPSFEVRKEKLGDRIAALQQLVSPFGKTDTA
SEQ ID NO:19  -LP---PSSKKPRLESHSSS----------MLPSFKVRKEKLGDRIAALQQLVSPFGKTDTA
SEQ ID NO:21  ------KKPRVESRSSC-------------PPFKVRKEKLGDRIAALQQLVSPFGKTDTA
                                                              MOTIF I

              ************************  ** **   *
SEQ ID NO:13  SVLMEAIGYIKFLQDQVETLSRPYLKSSRS-KKPRPTP-RGSSNASAGELEET--------
SEQ ID NO:16  SVLMEAIGYIKFLQDQVETLSRPYLKSSRSSKKPRPTHQRGCWNASAAGEEQEETTTR
SEQ ID NO:19  SVLMEAIGYIKFLQDQVETLSGPYLRSSKNSKKLACRAAQQRKGTSNGGDAAA-------
SEQ ID NO:21  SVLMEAIGYIKFLQSQIETLSVPYMRASRN------------RPGKASQLVSQSQEGDEEETR

              ************  ***            *    *
SEQ ID NO:13  --RPDLRSRGLCLVPLSCTSYVTNENGV-------WV-PPNFRA-N---------
SEQ ID NO:16  RRRPDLRSRGLCLVPLSCTSYVTNENGA-------WV-PSQF.-----------
SEQ ID NO:19  --KLDLRSRGLCLVPLSCTSYVTNENGV-------WP-PPNFRGN.--------
SEQ ID NO:21  ----DLRSRGLCLVPLSCMTYVTGDGGGDGGGGGVGTGFWPTPPGFGGGT.----
              MOTIF II
```

FIG. 4

% Identity

	SEQ ID NO:24 (B73-RT1)	SEQ ID NO:29 (cfp)	SEQ ID NO:39 (rive)	SEQ ID NO:25 (Arabido psis
SEQ ID NO:24 (B73-RT1)		47.2	39.7	26.0
SEQ ID NO:29 (cfp...)			33.2	24.1
SEQ ID NO:39 (rice)				29.1
SEQ ID NO:25 (Arabidopsist)				

FIG. 5

rrnB T2 transcription terminator
rrnB T1 transcription terminator
M13 (-40) forward primer
M13 (-20) forward primer
attP1
ccdB
Cm(R)
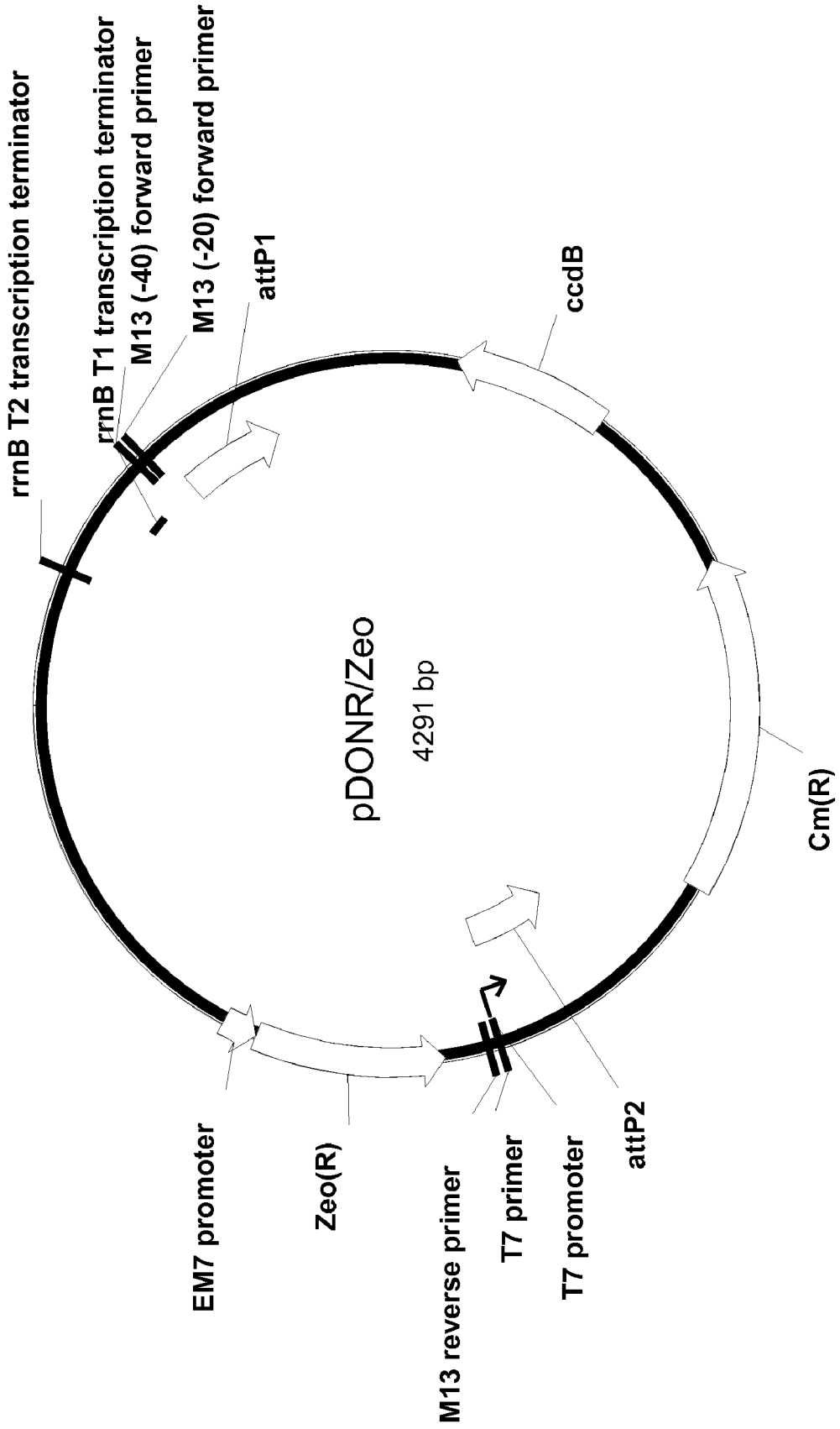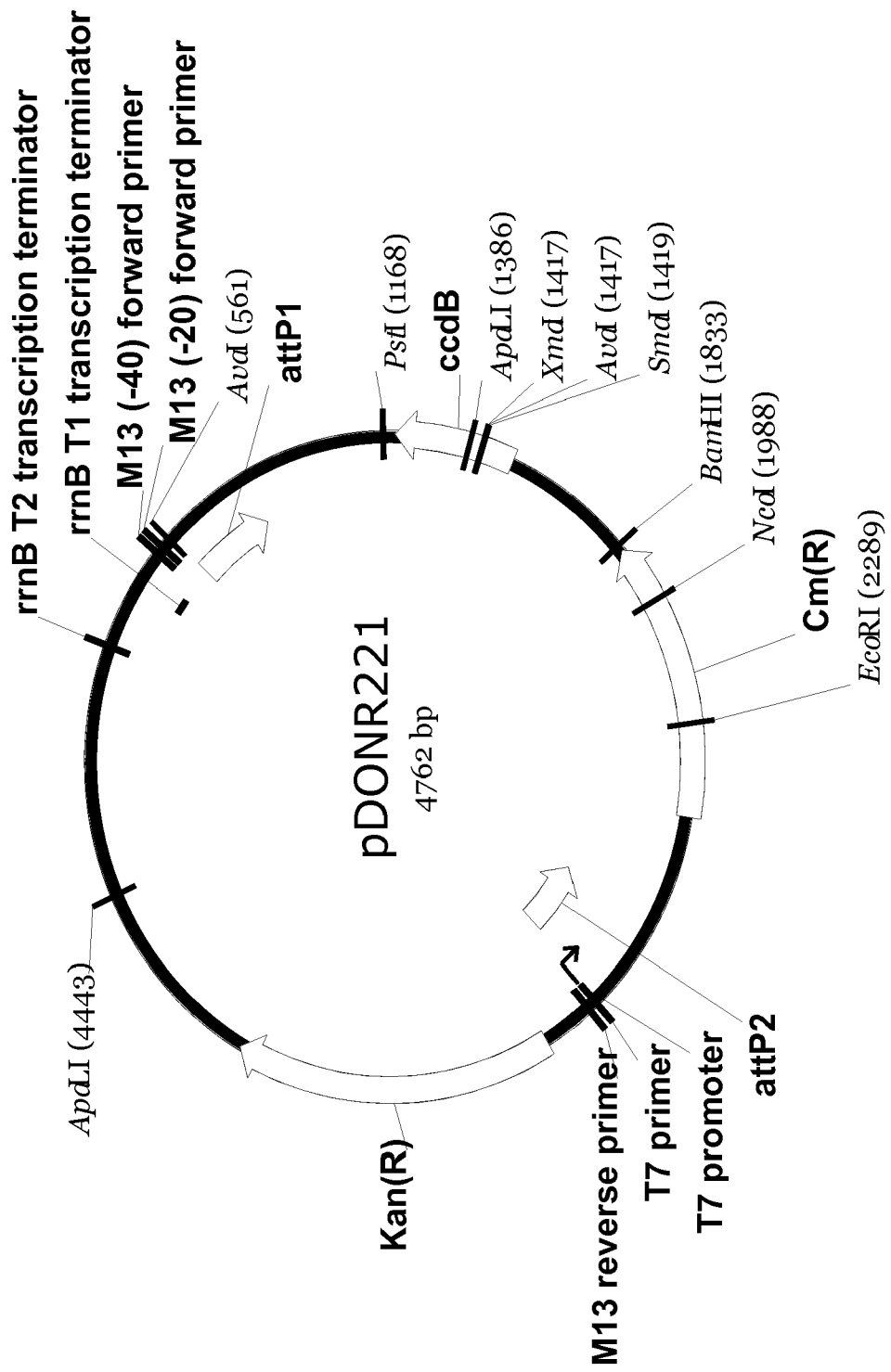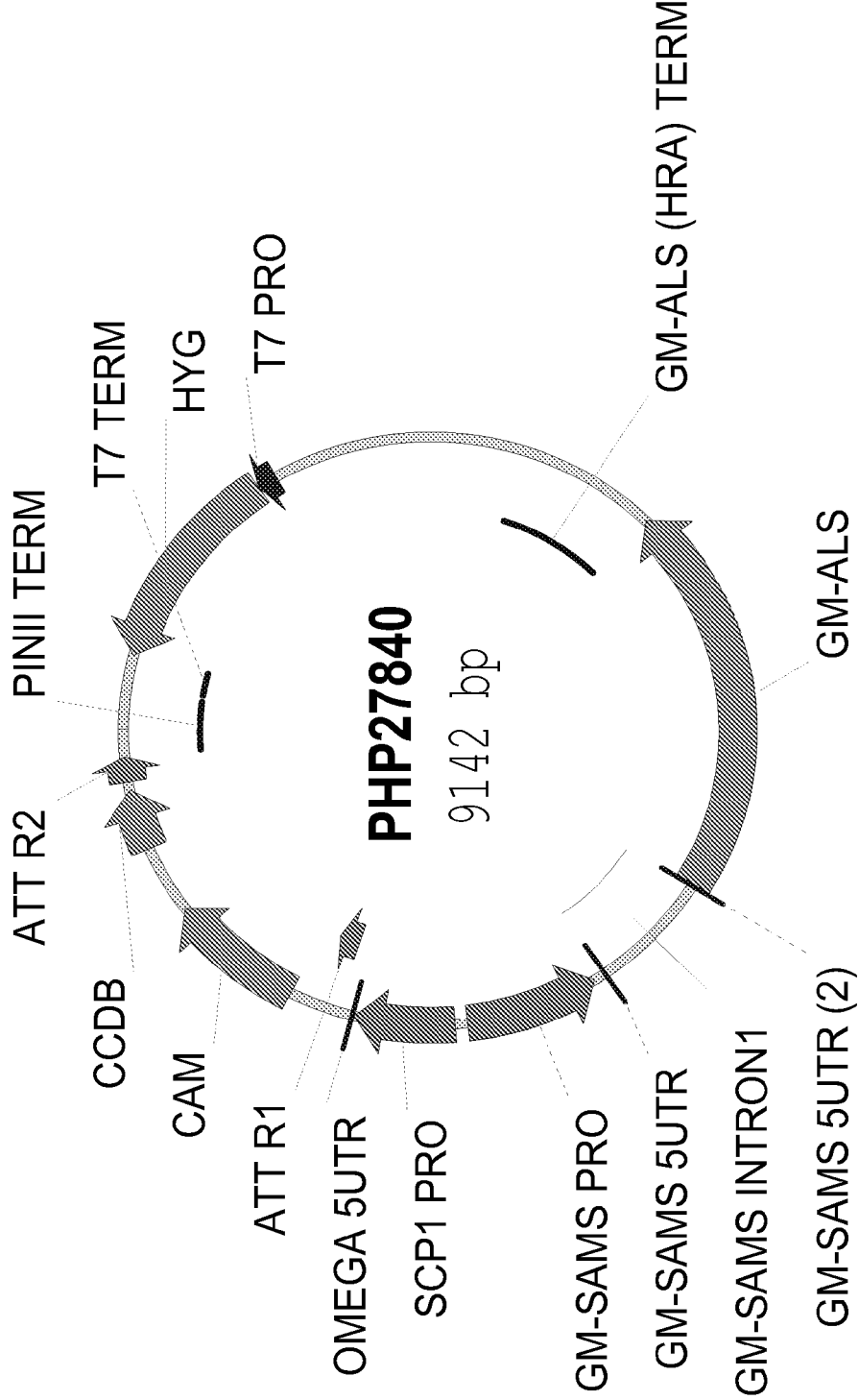attP2
T7 promoter
T7 primer
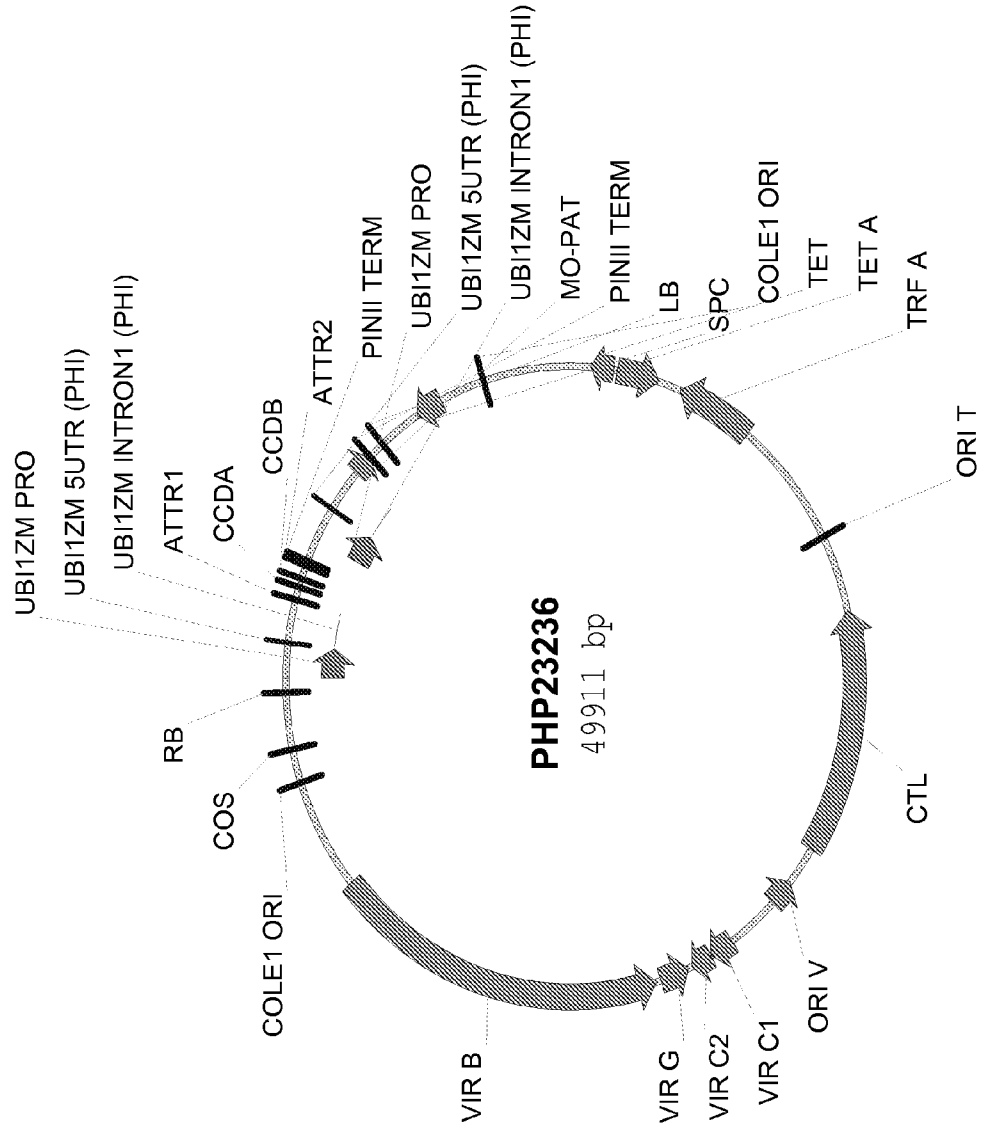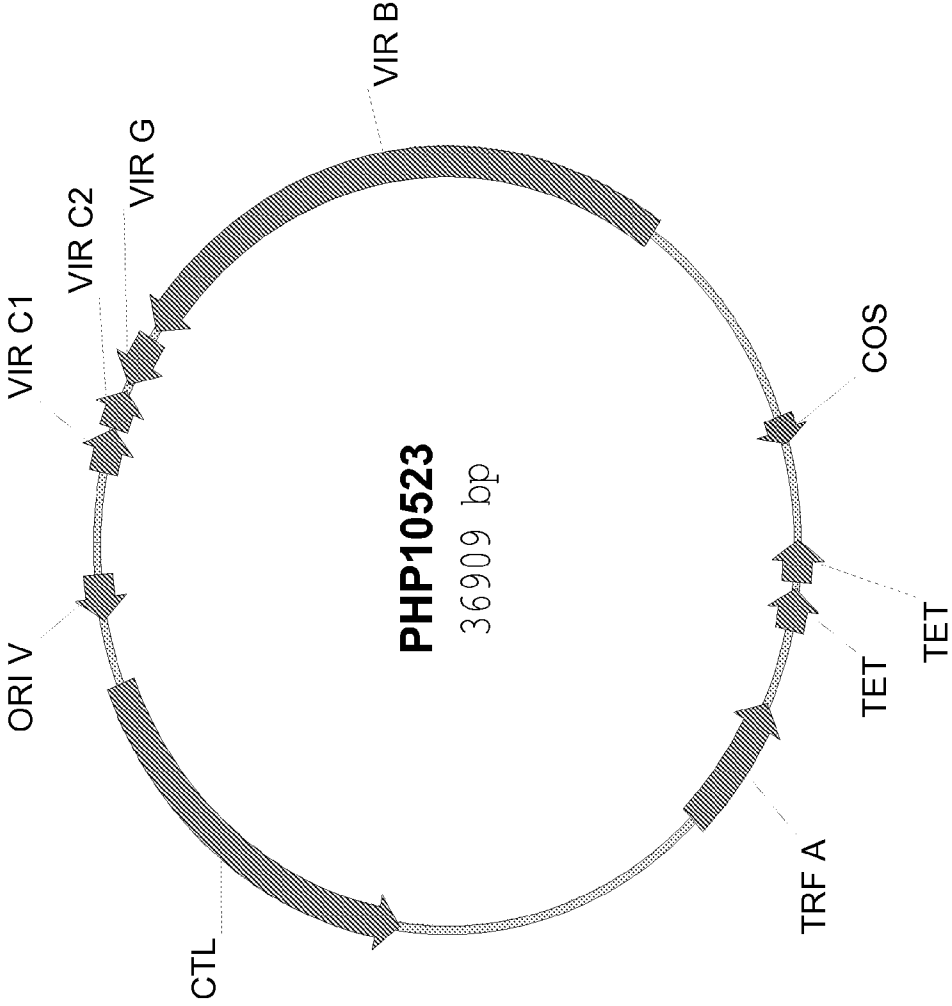M13 reverse primer
Zeo(R)
EM7 promoter

pDONR/Zeo
4291 bp

FIG. 6

FIG. 7

FIG. 8

FIG. 9
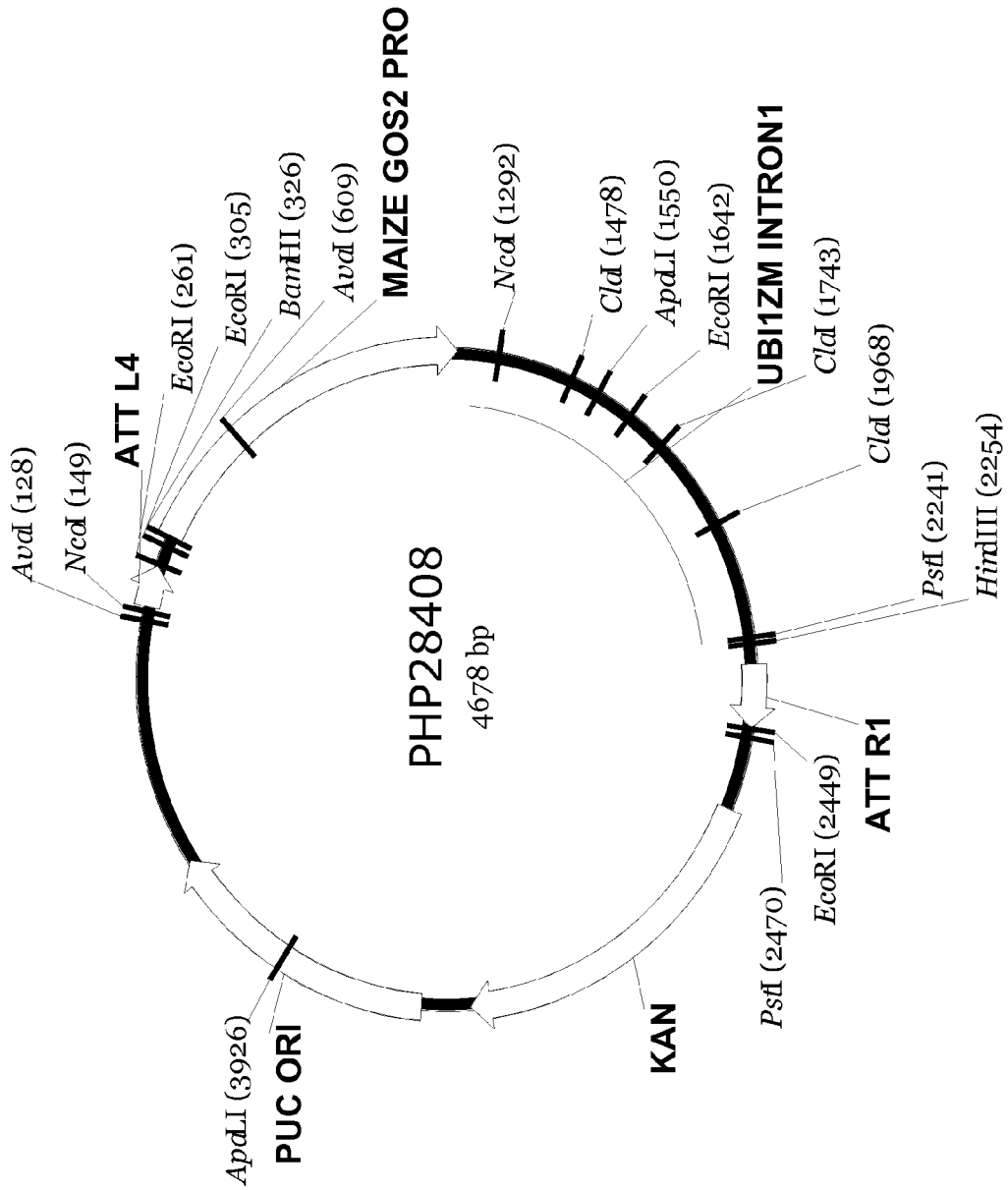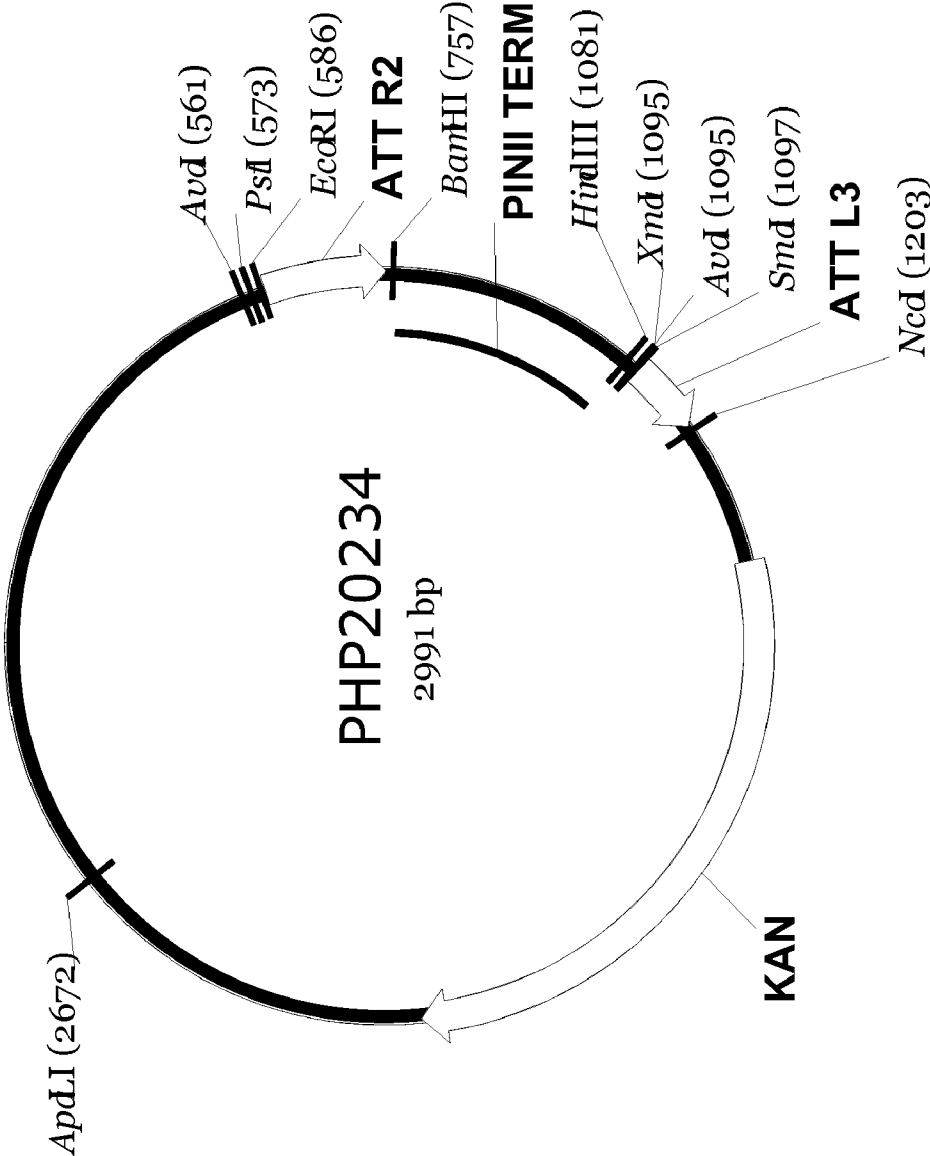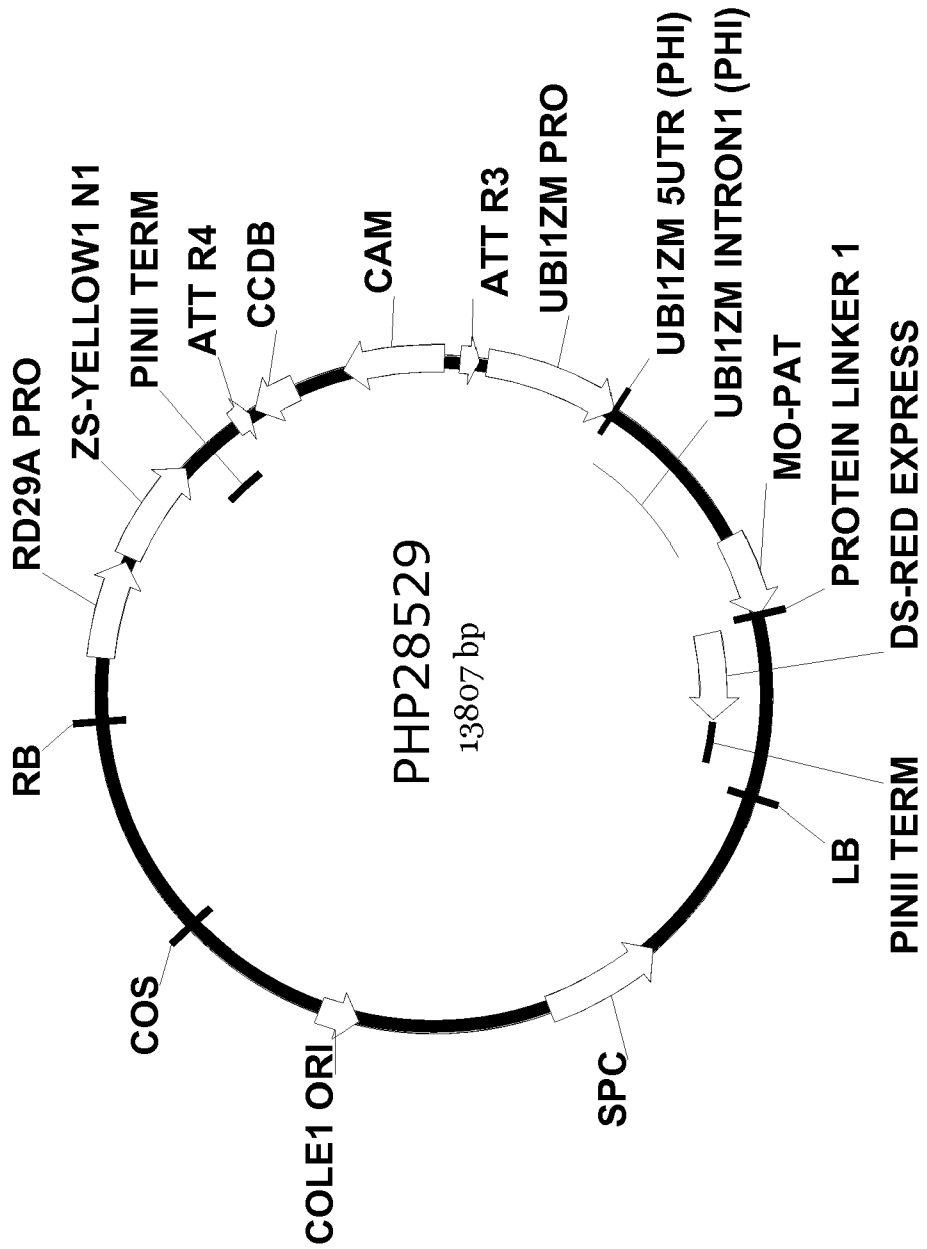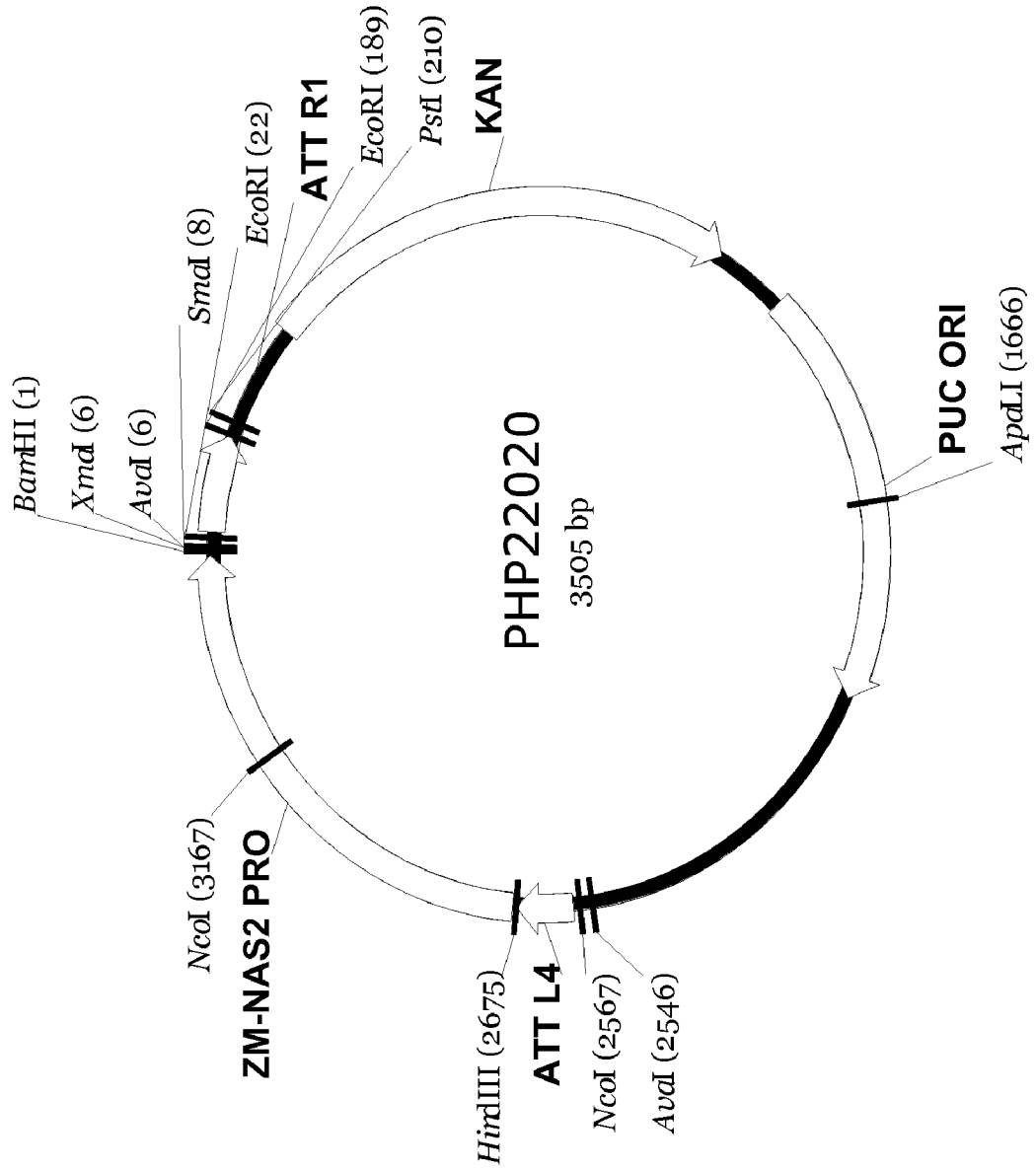
FIG. 10

FIG. 11

FIG. 12

FIG. 13

FIG. 14

PHP23112
4778 bp

Avd (128)
Ncd (149)
ATTL4
EcoRI (261)
EcoRI (305)
PsI (346)
UBI1ZM PRO
Avd (1037)
UBI1ZM 5UTR (PHI)
Cld (1573)
ApdI (1645)
EcoRI (1737)
UBI1ZM INTRON (PHI)
Cld (1838)
Cld (2063)
PsI (2336)
PsI (2570)
EcoRI (2549)
ATTR1
KAN
ApdI (4026)
PUC ORI

FIG. 15

FIG. 16

FIG. 17

PIIOXS2a-FRT87(ni)m
12856 bp
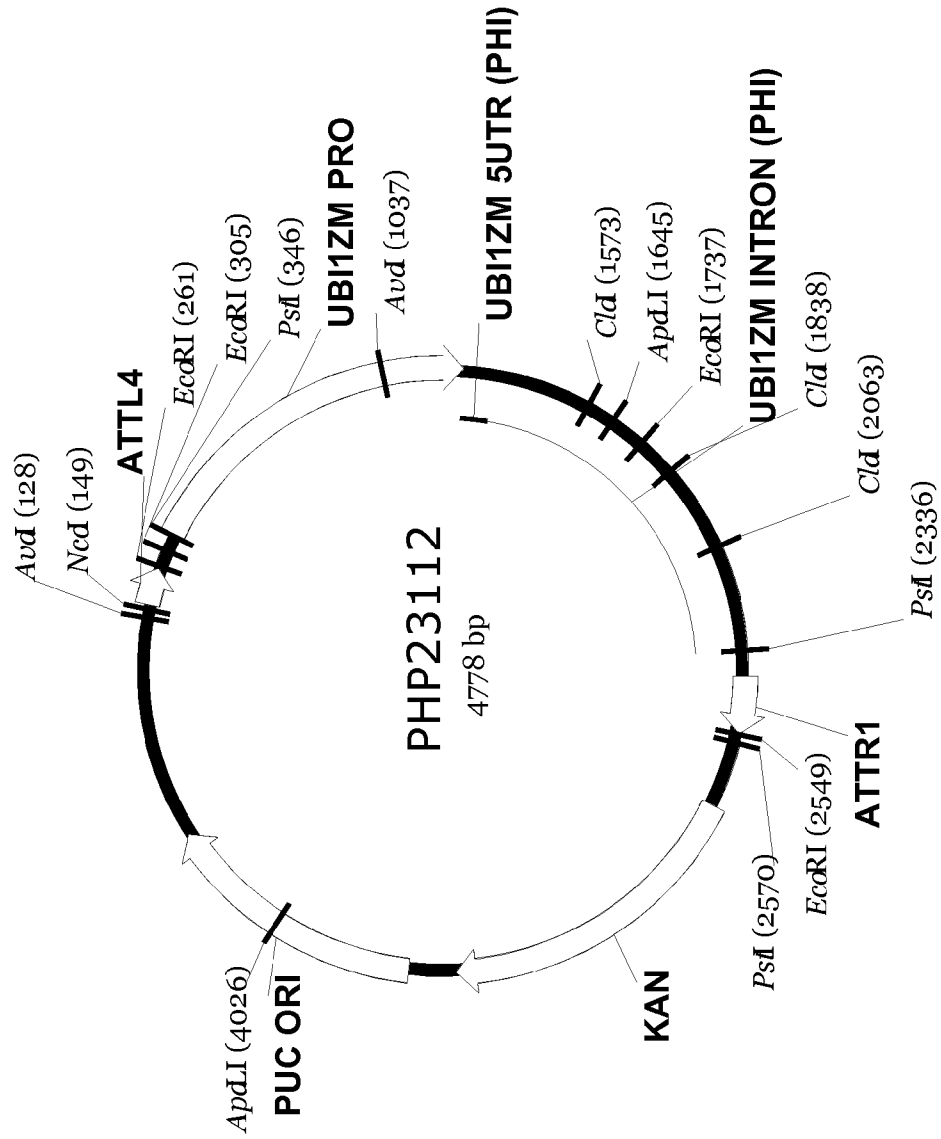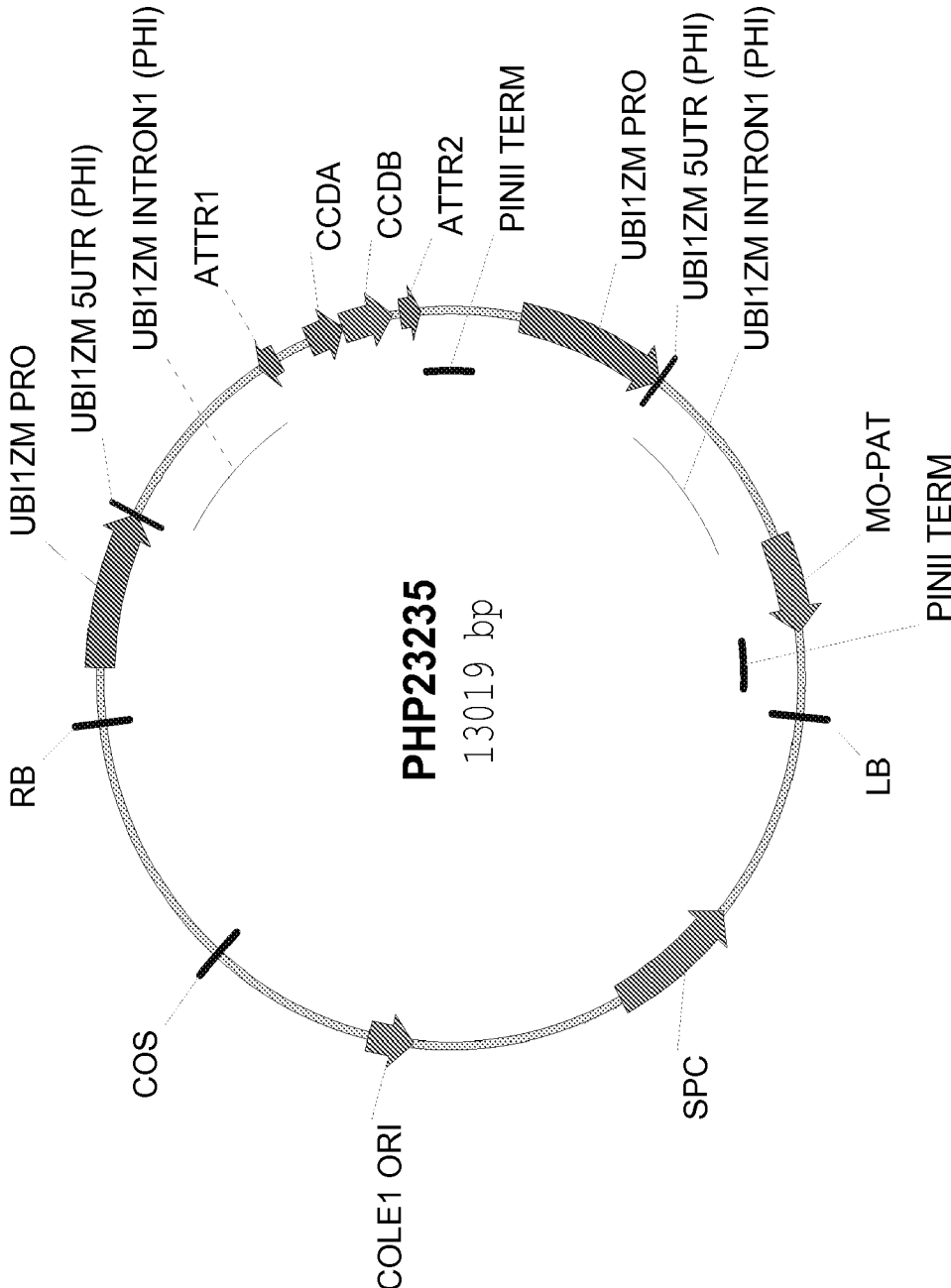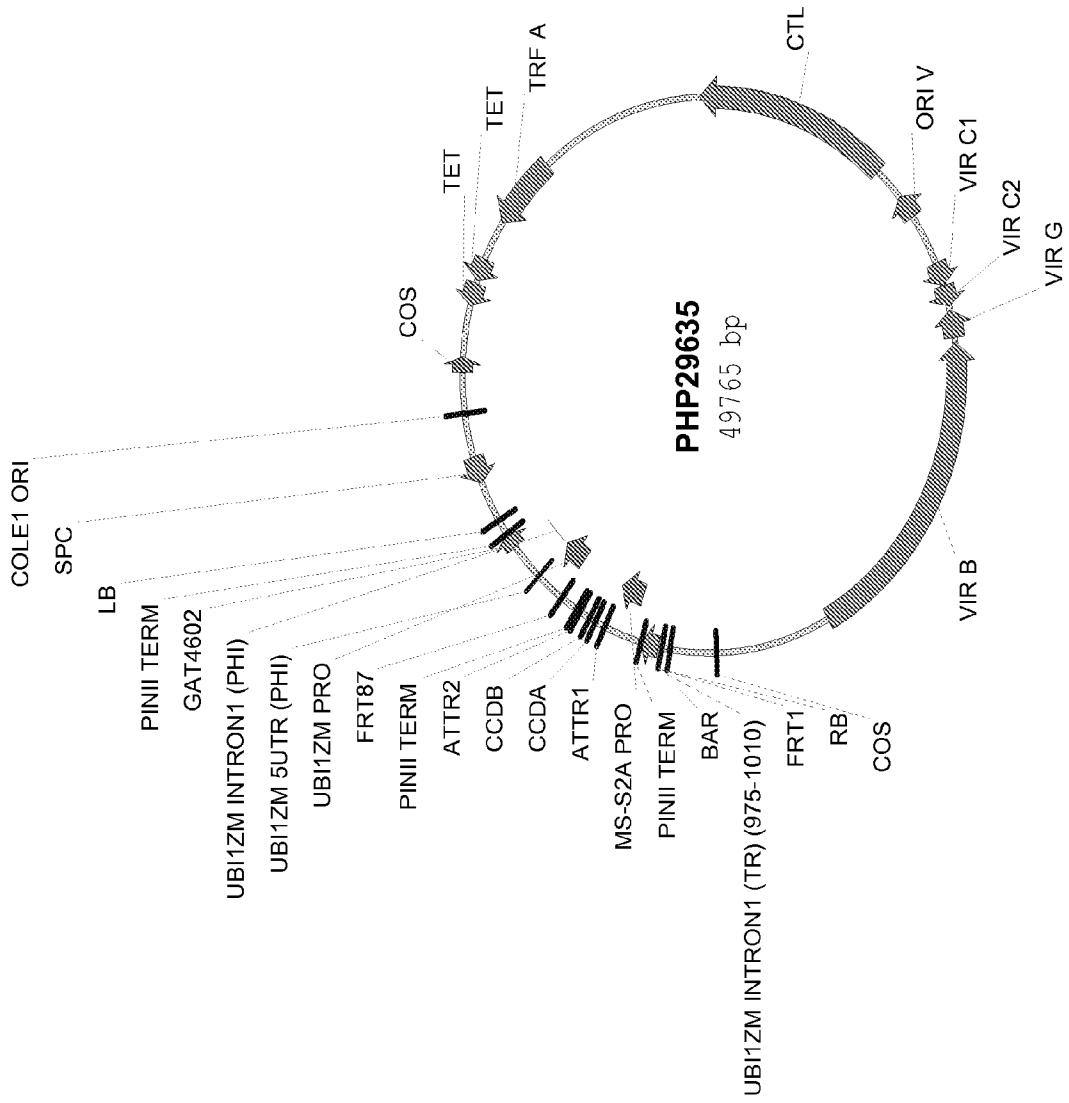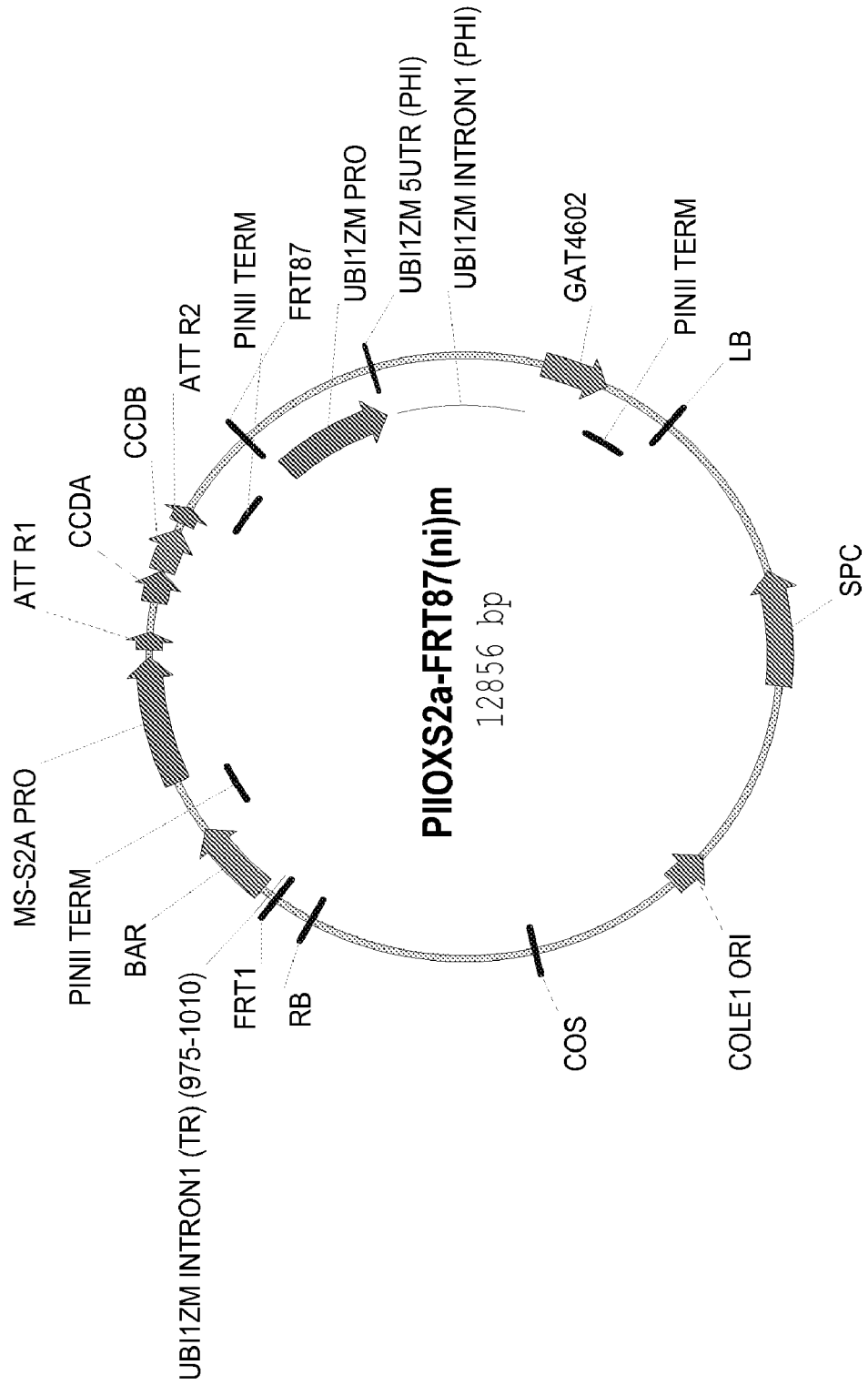
FIG. 18

Modified Hoagland's solutions -
16X concentrations for semi-hydroponics maize growth.

Nutrient	1 mM KNO_3	2 mM KNO_3	3 mM KNO_3	4 mM KNO_3
KNO_3	16 mM	32 mM	48 mM	64 mM
KCl	48 mM	32 mM	16 mM	------
KH_2PO_4	11 mM	11 mM	11 mM	11 mM
$MgSO_4$	16 mM	16 mM	16 mM	16 mM
$CaCl_2 \cdot 2H_2O$	16 mM	16 mM	16 mM	16 mM
Sprint 330	1.6 g/L	1.6 g/L	1.6 g/L	1.6 g/L
H_3BO_3	24 μM	24 μM	24 μM	24 μM
5 mM $MnCl_2 \cdot 4H_2O$	8 μM	8 μM	8 μM	8 μM
5 mM $ZnSO_4 \cdot 7 H_2O$	8 M	8 μM	8 μM	8 μM
0.5 mM $CuSO_4 \cdot 5 H_2O$	800 nM	800 nM	800 nM	800 nM
0.5 mM $H_2MoO_4 \cdot H_2O$	800 nM	800 nM	800 nM	800 nM

Dilute 16X with tap water and determine the pH of the final mixture.
Add 3-12 mL H_2SO_4 if the pH is above 6.5.
Optimum pH is 5.0 - 5.5

FIG. 19

The effect of different nitrate concentrations on the growth and development of Gaspe Bay Flint derived maize lines.

| [nitrate]	root (g dwt)	shoot (g dwt)	total vegetative (g dwt)	ear & husk (g dwt)	tassel (g dwt)	tiller #	tiller (g dwt)
1 week after emergence							
1 mM	0.070a	0.105b	0.175b				
2 mM	0.073a	0.137ab	0.209ab				
3 mM	0.056a	0.120ab	0.176ab				
4 mM	0.074a	0.157a	0.231a				
2 weeks after emergence							
1 mM	0.331ab	0.544c	0.875c				
2 mM	0.266b	0.951b	1.217b				
3 mM	0.352a	1.171a	1.523a				
4 mM	0.303ab	1.209a	1.512a				
3 weeks after emergence							
1 mM	0.757a	1.283b	2.040b	0.379c	0.239c	0.8c	0.080b
2 mM	0.785a	2.033a	2.819a	0.718a	0.363bc	2.3	0.506a
3 mM	0.664a	1.911a	2.574a	0.451bc	0.403ab	2.8ab	0.441a
4 mM	0.845a	2.129a	2.974a	0.650ab	0.506a	3.3a	0.688a
4 weeks after emergence							
1 mM	0.842b	2.010b	2.852b	1.318b	0.677b	*	*
2 mM	1.493a	3.772a	5.265a	3.130a	1.018a	*	*
3 mM	1.232ab	3.563a	4.795a	3.060a	0.875ab	*	*
4 mM	1.010b	2.943a	3.952a	2.787a	0.891ab	*	*

* Tillers removed 3 weeks after emergence
Means with similar letters are not different by protected Least Significant Difference (LSD) (0.05)

FIG. 20a

rt1 Phenotypic Characterization
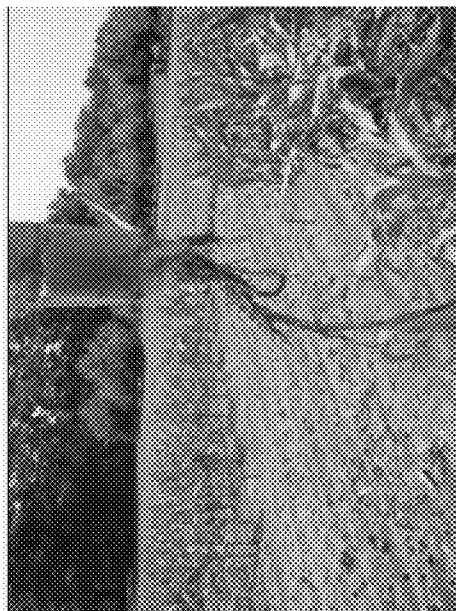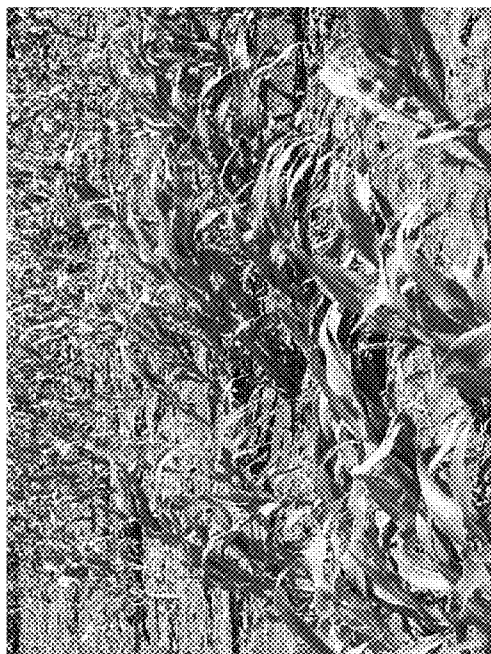
rt1 in greenhouse

rt1 in field

FIG. 20b
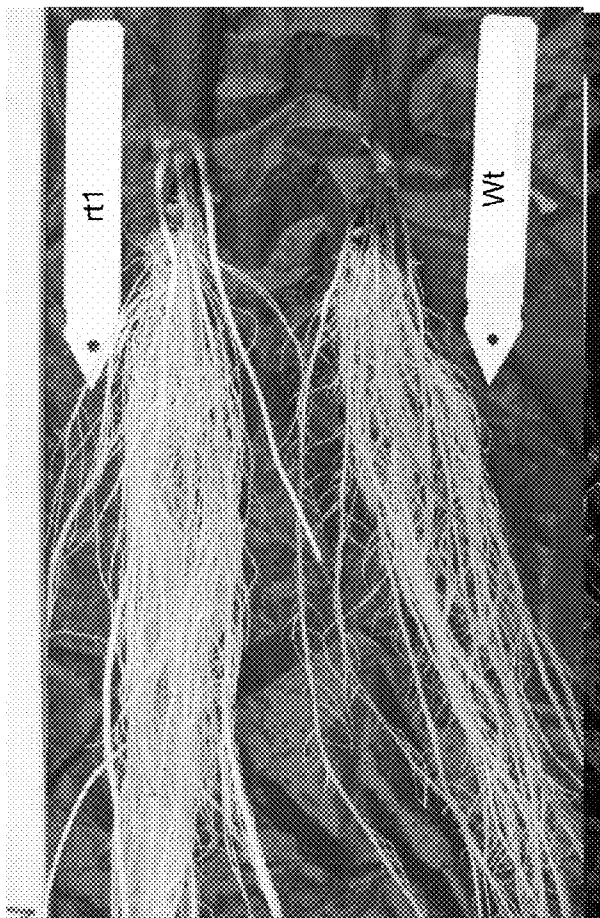
In a hydroponic system

FIG. 20c

In autoclaved hard
soil in the
Greenhouse

Left: Wt plants

Right: rt1 mutant

PLANTS WITH ALTERED ROOT ARCHITECTURE, INVOLVING THE RT1 GENE, RELATED CONSTRUCTS AND METHODS

FIELD OF THE INVENTION

[0001]   This invention relates to compositions and methods useful in altering root architecture in plants. Additionally, the invention relates to plants that have been genetically transformed with the compositions of the invention.

BACKGROUND OF THE INVENTION

[0002]   Relatively little is known about the genetic regulation of plant root development and function. Elucidation of the genetic regulation is important because roots serve important functions such as acquisition of water and nutrients and the anchorage of the plants in the soil.

[0003]   Maize root architecture is composed of different root types formed at different plant developmental stages. A number of mutants affected in specific root types during different developmental stages have been described in maize (e.g. rtcs (rootless concerning crown and seminal roots), lrt1 (lateral rootless1)), rt1 ((rootless 1) (Hochholdinger et al. (2004) Annals of Botany 93: 359-368). The mutant rt1 was the first mutant of root formation that was isolated and shows a reduced number of shoot-borne roots. The rt1 mutant is missing all shoot-borne roots at the higher nodes while there is only a slight difference in the number of crown roots at the first two nodes. The mutation rt1 is inherited as a monogenic recessive trait and maps on chromosome 3 (Maize GDB on the World Wide Web at maizegdb.org)

[0004]   The rt1 mutant was first described by Jenkins (Jenkins M T (1930) J Hered 21:79-80), but there, has been no molecular analysis of the nucleic acid encoding the protein associated with the rt1 phenotype. Indeed, the identity of the protein encoded by rt1 has not been reported, so far.

SUMMARY OF THE INVENTION

[0005]   The present invention includes:

[0006]   In one embodiment an isolated polynucleotide comprising: (i) a nucleic acid sequence encoding a polypeptide having an amino acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13 or 21 and wherein expression of said polypeptide in a plant results in an altered root architecture when compared to a control plant not comprising said recombinant DNA construct, or (ii) a full complement of the nucleic acid sequence of (i), wherein the full complement and the nucleic acid sequence of (i) consist of the same number of nucleotides and are 100% complementary.

[0007]   In another embodiment an isolated polypeptide having an amino acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13 or 21 and wherein expression of said polypeptide in a plant results in an altered plant root architecture when compared to a control plant not comprising said recombinant DNA construct.

[0008]   In yet another embodiment an isolated polynucleotide comprising (i) a nucleic acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 12 or 20 and wherein said polynucleotide encodes a polypeptide wherein expres-

sion of said polypeptide results in an altered root architecture when compared to a control plant not comprising said recombinant DNA construct or (ii) a full complement of the nucleic acid sequence of (i).

[0009]   In yet another embodiment an isolated polynucleotide comprising (i) a nucleic acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 12 or 20 and wherein said polynucleotide encodes a polypeptide wherein expression of said polypeptide results in an altered root architecture when compared to a control plant not comprising said recombinant DNA construct and wherein the polypeptide sequence comprises at least one motif selected from the group consisting of SEQ ID NOs:22 and 23, wherein said motif is a substantially conserved subsequence.

[0010]   In another embodiment an isolated polynucleotide encoding a polypeptide, wherein expression of said polypeptide results in an altered root architecture and wherein the polypeptide sequence comprises at least one motif selected from the group consisting of SEQ ID NOs:22 and 23, wherein said motif is a substantially conserved subsequence.

[0011]   In still another embodiment, a plant comprising in its genome a recombinant DNA construct comprising a polynucleotide operably linked to at least one regulatory element, wherein said polynucleotide encodes a polypeptide having an amino acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13, 17, 19, or 21, and wherein said plant exhibits altered root architecture when compared to a control plant not comprising said recombinant DNA construct.

[0012]   In another embodiment, a plant comprising in its genome a recombinant DNA construct comprising:

[0013]   (a) a polynucleotide operably linked to at least one regulatory element, wherein said polynucleotide encodes a polypeptide having an amino acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13, 17, 19, or 21, or

[0014]   (b) a suppression DNA construct comprising at least one regulatory element operably linked to: (i) all or part of: (A) a nucleic acid sequence encoding a polypeptide having an amino acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13, 17, 19, or 21, or (B) a full complement of the nucleic acid sequence of (b)(i)(A); or (ii) a region derived from all or part of a sense strand or antisense strand of a target gene of interest, said region having a nucleic acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to said all or part of a sense strand or antisense strand from which said region is derived, and wherein said target gene of interest encodes a RT1 or RT1-like polypeptide, and wherein said plant exhibits an alteration of at least one agronomic characteristic when compared to a control plant not comprising said recombinant DNA construct.

[0015]   In another embodiment, a method of altering root architecture in a plant, comprising (a) introducing into a regenerable plant cell a recombinant DNA construct comprising a polynucleotide operably linked to at least one regulatory sequence, wherein the polynucleotide encodes a polypeptide having an amino acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13, 17, 19, or 21; and (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome

the recombinant DNA construct and exhibits altered root architecture when compared to a control plant not comprising the recombinant DNA construct; and optionally, (c) obtaining a progeny plant derived from the transgenic plant, wherein said progeny plant comprises in its genome the recombinant DNA construct and exhibits altered root architecture when compared to a control plant not comprising the recombinant DNA construct.

[0016] In another embodiment, a method of evaluating root architecture in a plant, comprising (a) introducing into a regenerable plant cell a recombinant DNA construct comprising a polynucleotide operably linked to at least one regulatory sequence, wherein the polynucleotide encodes a polypeptide having an amino acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13, 17, 19, or 21; (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the recombinant DNA construct; and (c) evaluating root architecture of the transgenic plant compared to a control plant not comprising the recombinant DNA construct; and optionally, (d) obtaining a progeny plant derived from the transgenic plant, wherein the progeny plant comprises in its genome the recombinant DNA construct; and optionally, (e) evaluating root architecture of the progeny plant compared to a control plant not comprising the recombinant DNA construct.

[0017] In another embodiment, a method of evaluating root architecture in a plant, comprising (a) introducing into a regenerable plant cell a recombinant DNA construct comprising a polynucleotide operably linked to at least one regulatory sequence, wherein the polynucleotide encodes a polypeptide having an amino acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13, 17, 19, or 21; (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the recombinant DNA construct; (c) obtaining a progeny plant derived from the transgenic plant, wherein the progeny plant comprises in its genome the recombinant DNA construct; and (d) evaluating root architecture of the progeny plant compared to a control plant not comprising the recombinant DNA construct.

[0018] In another embodiment, a method of determining an alteration of an agronomic characteristic in a plant, comprising (a) introducing into a regenerable plant cell a recombinant DNA construct comprising a polynucleotide operably linked to at least one regulatory sequence, wherein the polynucleotide encodes a polypeptide having an amino acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13, 17, 19, or 21; (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the recombinant DNA construct; and (c) determining whether the transgenic plant exhibits an alteration of at least one agronomic characteristic when compared to a control plant not comprising the recombinant DNA construct; and optionally, (d) obtaining a progeny plant derived from the transgenic plant, wherein the progeny plant comprises in its genome the recombinant DNA construct; and optionally, (e) determining whether the progeny plant exhibits an alteration of at least one agronomic characteristic when compared to a control plant not comprising the recombinant DNA construct.

[0019] In another embodiment, a method of determining an alteration of an agronomic characteristic in a plant, comprising (a) introducing into a regenerable plant cell a recombinant DNA construct comprising a polynucleotide operably linked to at least one regulatory sequence, wherein the polynucleotide encodes a polypeptide having an amino acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13, 17, 19, or 21; (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the recombinant DNA construct; (c) obtaining a progeny plant derived from the transgenic plant, wherein the progeny plant comprises in its genome the recombinant DNA construct; and (d) determining whether the progeny plant exhibits an alteration of at least one agronomic characteristic when compared to a control plant not comprising the recombinant DNA construct.

[0020] In another embodiment, a method of determining an alteration of an agronomic characteristic in a plant, comprising:

[0021] (a) introducing into a regenerable plant cell a suppression DNA construct comprising at least one regulatory element operably linked to:

[0022] (i) all or part of: (A) a nucleic acid sequence encoding a polypeptide having an amino acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13, 17, 19, or 21, or (B) a full complement of the nucleic acid sequence of (b)(i)(A); or

[0023] (ii) a region derived from all or part of a sense strand or antisense strand of a target gene of interest, said region having a nucleic acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to said all or part of a sense strand or antisense strand from which said region is derived, and wherein said target gene of interest encodes a RT1 or RT1-like polypeptide;

[0024] (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the suppression DNA construct; and

[0025] (c) determining whether the transgenic plant exhibits an alteration of at least one agronomic characteristic when compared to a control plant not comprising the suppression DNA construct;

[0026] and optionally, (d) obtaining a progeny plant derived from the transgenic plant, wherein the progeny plant comprises in its genome the suppression DNA construct; and optionally, (e) determining whether the progeny plant exhibits an alteration of at least one agronomic characteristic when compared to a control plant not comprising the suppression DNA construct.

[0027] In another embodiment, a method of determining an alteration of an agronomic characteristic in a plant, comprising:

[0028] (a) introducing into a regenerable plant cell a suppression DNA construct comprising at least one regulatory element operably linked to:

[0029] (i) all or part of: (A) a nucleic acid sequence encoding a polypeptide having an amino acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13, 17, 19, or 21, or (B) a full complement of the nucleic acid sequence of (a)(i)(A); or

[0030] (ii) a region derived from all or part of a sense strand or antisense strand of a target gene of interest, said region having a nucleic acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to said all or part of a sense strand or antisense strand from which said region is derived, and wherein said target gene of interest encodes a RT1 or RT1-like polypeptide;

[0031] (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the suppression DNA construct and exhibits altered root architecture when compared to a control plant not comprising the suppression DNA construct;

[0032] (c) obtaining a progeny plant derived from the transgenic plant, wherein the progeny plant comprises in its genome the suppression DNA construct; and

[0033] (d) determining whether the progeny plant exhibits an alteration of at least one agronomic characteristic when compared to a control plant not comprising the suppression DNA construct.

[0034] In another embodiment, a method of altering root architecture in a plant, comprising:

[0035] (a) introducing into a regenerable plant cell a suppression DNA construct comprising at least one regulatory element operably linked to:

[0036] (i) all or part of: (A) a nucleic acid sequence encoding a polypeptide having an amino acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13, 17, 19, or 21; or (B) a full complement of the nucleic acid sequence of (a)(i)(A); or

[0037] (ii) a region derived from all or part of a sense strand or antisense strand of a target gene of interest, said region having a nucleic acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to said all or part of a sense strand or antisense strand from which said region is derived, and wherein said target gene of interest encodes a RT1 or RT1-like polypeptide; and

[0038] (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the suppression DNA construct and wherein the transgenic plant exhibits altered root architecture when compared to a control plant not comprising the suppression DNA construct; and optionally, (c) obtaining a progeny plant derived from the transgenic plant, wherein said progeny plant comprises in its genome the recombinant DNA construct and wherein the progeny plant exhibits altered root architecture when compared to a control plant not comprising the suppression DNA construct.

[0039] In another embodiment, a method of evaluating root architecture in a plant, comprising:

[0040] (a) introducing into a regenerable plant cell a suppression DNA construct comprising at least one regulatory element operably linked to:

[0041] (i) all or part of: (A) a nucleic acid sequence encoding a polypeptide having an amino acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13, 17, 19, or 21; or (B) a full complement of the nucleic acid sequence of (a)(i)(A); or

[0042] (ii) a region derived from all or part of a sense strand or antisense strand of a target gene of interest, said region having a nucleic acid sequence of at least 50%

sequence identity, based on the Clustal V method of alignment, when compared to said all or part of a sense strand or antisense strand from which said region is derived, and wherein said target gene of interest encodes a RT1 or RT1-like polypeptide;

[0043] (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the suppression DNA construct; and

[0044] (c) evaluating root architecture of the transgenic plant compared to a control plant not comprising the suppression DNA construct;

[0045] and optionally, (d) obtaining a progeny plant derived from the transgenic plant, wherein the progeny plant comprises in its genome the suppression DNA construct; and optionally, (e) evaluating root architecture of the progeny plant compared to a control plant not comprising the suppression DNA construct.

[0046] In another embodiment, a method of evaluating root architecture in a plant, comprising:

[0047] (a) introducing into a regenerable plant cell a suppression DNA construct comprising at least one regulatory element operably linked to:

[0048] (i) all or part of: (A) a nucleic acid sequence encoding a polypeptide having an amino acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13, 17, 19, or 21, or (B) a full complement of the nucleic acid sequence of (a)(i)(A); or

[0049] (ii) a region derived from all or part of a sense strand or antisense strand of a target gene of interest, said region having a nucleic acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to said all or part of a sense strand or antisense strand from which said region is derived, and wherein said target gene of interest encodes a RT1 or RT1-like polypeptide;

[0050] (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the suppression DNA construct;

[0051] (c) obtaining a progeny plant derived from the transgenic plant, wherein the progeny plant comprises in its genome the suppression DNA construct; and

[0052] (d) evaluating root architecture of the progeny plant compared to a control plant not comprising the suppression DNA construct.

[0053] Also included in the present invention is any progeny of the above plants, any seeds of the above plants, and cells from any of the above plants and progeny.

[0054] A method of producing seed that can be sold as a product offering with altered root architecture comprising any of the preceding preferred methods, and further comprising obtaining seeds from said progeny plant, wherein said seeds comprise in their genome said recombinant DNA construct.

BRIEF DESCRIPTION OF THE FIGURES AND SEQUENCE LISTINGS

[0055] The invention can be more fully understood from the following detailed description and the accompanying drawings and Sequence Listing which form a part of this application.

[0056] FIG. 1 depicts the RT1 gene on bac clone b0541.c13 (SEQ ID NO:9)

4

[0057] FIG. 2 demonstrates induction of RT1 transcripts after addition of Ethephon.

[0058] FIGS. 3A-3B show the multiple alignment of the full length amino acid sequences of RT1 protein from B73 (SEQ ID NO:13), the rice RT1 homolog (NCBI General identifier No. SEQ ID NO:17), the *Arabidopsis* RT1 homolog (NCBI General identifier No. SEQ ID NO:19) and the maize RT1 homolog from clone cfp7n.pk6.13 (SEQ ID NO:21). Amino acids conserved among all sequences are indicated with an asterisk (*) on the top row; dashes are used by the program to maximize alignment of the sequences. Two highly conserved motifs among all four sequences are underlined in the alignment. The method parameters used to produce the multiple alignment of the sequences below was performed using the Clustal method of alignment (Higgins and Sharp (1989) *CABIOS.* 5:151-153) with the default parameters (GAP PENALTY=10, GAP LENGTH PENALTY=10).

[0059] FIG. 4 shows a chart of the percent sequence identity for each pair of amino acid sequences displayed in FIGS. 3A-3B.

[0060] FIG. 5 depicts the vector pDONOR™/Zeo.

[0061] FIG. 6 depicts the vector pDONOR™221.

[0062] FIG. 7 depicts the vector PHP27840.

[0063] FIG. 8 depicts the vector PHP23236.

[0064] FIG. 9 depicts the vector PHP10523.

[0065] FIG. 10 depicts the vector PHP28408.

[0066] FIG. 11 depicts the vector PHP20234.

[0067] FIG. 12 depicts the vector PHP28529.

[0068] FIG. 13 depicts the vector PHP22020.

[0069] FIG. 14 depicts the vector PHP23112.

[0070] FIG. 15 depicts the vector PHP23235.

[0071] FIG. 16 depicts the vector PHP29635.

[0072] FIG. 17 depicts the vector pIIOXS2a-FRT87(ni)m.

[0073] FIG. 18 is the growth medium used for semi-hydroponics maize growth in Example 19.

[0074] FIG. 19 is a chart setting forth data relating to the effect of different nitrate concentrations on the growth and development of Gaspe Bay Flint derived maize lines in Example 19.

[0075] FIG. 20 *a-c* show a comparison of rt1 and wild type plants grown in the field, greenhouse or hydroponic conditions.

[0076] The sequence descriptions and Sequence Listing attached hereto comply with the rules governing nucleotide and/or amino acid sequence disclosures in patent applications as set forth in 37 C.F.R. §1.821-1.825.

[0077] The Sequence Listing contains the one letter code for nucleotide sequence characters and the three letter codes for amino acids as defined in conformity with the IUPAC-IUBMB standards described in *Nucleic Acids Res.* 13:3021-3030 (1985) and in the *Biochemical J.* 219 (No. 2):345-373 (1984) which are herein incorporated by reference. The symbols and format used for nucleotide and amino acid sequence data comply with the rules set forth in 37 C.F.R. §1.822.

[0078] SEQ ID NO:1 is the forward primer for marker MZA8757-F81 used in Example 2.

[0079] SEQ ID NO:2 is the reverse primer for marker MZA8757-R593 used in Example 2.

[0080] SEQ ID NO:3 is the forward primer for marker MZA15417-F132 used in Example 2.

[0081] SEQ ID NO:4 is the reverse primer for marker MZA15417-R607 used in Example 1.

[0082] SEQ ID NO:5 is the forward primer for CAP marker b0541 used in Example 2.

[0083] SEQ ID NO:6 is the reverse primer for CAP marker b0541 used in Example 2.

[0084] SEQ ID NO:7 is the forward primer for CAP marker b0461 used in Example 2

[0085] SEQ ID NO:8 is the reverse primer for CAP marker b0461 used in Example 2.

[0086] SEQ ID NO:9 is the sequence of the candidate gene derived from BAC b0541.c13 described in Example 3.

[0087] SEQ ID NO:10 is the forward primer RT1 3006F used in Example 3.

[0088] SEQ ID NO:11 is the reverse primer RT1 17631R used in Example 3.

[0089] SEQ ID NO:12 is the B73 cDNA of RT1 described in Example 3.

[0090] SEQ ID NO:13 is the RT1 amino acid sequence encoded by nucleotides 50 through 1382 (Stop) of SEQ ID NO:12.

[0091] SEQ ID NO:14 is the forward primer 4405F used in Example 5.

[0092] SEQ ID NO:15 is the reverse primer etr4Rnew used in Example 5.

[0093] SEQ ID NO:16 is the nucleotide sequence encoding the closest polypeptide RT1 homolog from rice.

[0094] SEQ ID NO:17 corresponds to the RT1 amino acid sequence homolog encoded by nucleotides 91 through 1263 (Stop) of SEQ ID NO:16 and is set forth in NCBI General identifier No. 115434026.

[0095] SEQ ID NO:18 is the nucleotide sequences encoding the closest polypeptide RT1 homolog from *Arabidopsis*.

[0096] SEQ ID NO:19 corresponds to the RT1 amino acid sequence homolog encoded by nucleotides 132 through 1493 (Stop) of SEQ ID NO:18 and is set forth in NCBI General identifier No. 15217667.

[0097] SEQ ID NO:20 is an EST corresponding to a maize homolog of the maize RT1 sequence.

[0098] SEQ ID NO:21 is the amino acid sequence encoded by SEQ ID NO:20.

[0099] SEQ ID NO:22 corresponds to Motif I in the alignment shown in FIGS. 3A-3B.

[0100] SEQ ID NO:23 corresponds to Motif II in the alignment shown in FIGS. 3A-3B.

[0101] SEQ ID NO:24 is the attB1 sequence described in Example 9.

[0102] SEQ ID NO:25 is the attB2 sequence described in Example 9.

[0103] SEQ ID NO:26 is the sequence of the forward primer VC062 described in Example 9.

[0104] SEQ ID NO:27 is the sequence of the reverse primer VC063 described in Example 9.

[0105] SEQ ID NO:28 is the sequence of vector pDONOR™/Zeo described in Example 9.

[0106] SEQ ID NO:29 is the sequence of vector pDONOR™/221 described in Example 9.

[0107] SEQ ID NO:30 is the sequence of PHP27840 described in Example 9.

[0108] SEQ ID NO:31 is the sequence of PHP23236 described in Example 9.

[0109] SEQ ID NO:32 is the sequence of PHP10523.

[0110] SEQ ID NO:33 is the sequence of the NAS2 promoter.

[0111] SEQ ID NO:34 is the sequence of the GOS2 promoter.

[0112] SEQ ID NO:35 is the sequence of the ubiquitin promoter.

5

[0113] SEQ ID NO:36 is the sequence of the PINII termi-nator.

[0114] SEQ ID NO:37 is the sequence of PHP28408.

[0115] SEQ ID NO:38 is the sequence of PHP20234.

[0116] SEQ ID NO:39 is the sequence of PHP28529.

[0117] SEQ ID NO:40 is the sequence of PHP22020.

[0118] SEQ ID NO:41 is the sequence of PHP23112.

[0119] SEQ ID NO:42 is the sequence of PHP23235.

[0120] SEQ ID NO:43 is the sequence of PHP29635.

[0121] SEQ ID NO:44 is the sequence of pHOXS2a-FRT87 (ni)m.

[0122] SEQ ID NO:45 is the sequence of the S2A promoter.

DETAILED DESCRIPTION

[0123] The disclosure of each reference set forth herein is hereby incorporated by reference in its entirety.

[0124] As used herein and in the appended claims, the singular forms "a", "an", and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, reference to "a plant" includes a plurality of such plants, reference to "a cell" includes one or more cells and equivalents thereof known to those skilled in the art, and so forth.

[0125] The term "root architecture" refers to the arrange-ment of different plant parts that comprise the root. The terms "root architecture", "root structure", "root system" and "root system architecture" are used interchangeably herein.

[0126] In general, the first root of a plant that develops from the embryo is called the primary root. In most dicots, the primary root is called the taproot. This taproot grows down-ward and gives rise to branch (lateral) roots. In monocots the primary root of the plant branches, giving rise to a fibrous root system.

[0127] The term "altered root architecture" refers to changes in the different parts that make up the root system at different stages of its development compared to a reference or control plant. It is understood that altered root architecture encompasses changes in one or more measurable parameters, including and not limited to, the diameter, length, number, angle or surface of one or more of the root system parts, including and not limited to, the primary root, lateral or branch root, crown roots, adventitious root, and root hairs, all of which fall within the scope of this invention. These changes can lead to an overall alteration in the area or volume occu-pied by the root.

[0128] "Agronomic characteristics" is a measurable param-eter including and not limited to greenness, yield, growth rate, biomass, fresh weight at maturation, dry weight at matura-tion, fruit yield, seed yield, total plant nitrogen content, fruit nitrogen content, seed nitrogen content, nitrogen content in a vegetative tissue, total plant free amino acid content, fruit free amino acid content, seed free amino acid content, free amino acid content in a vegetative tissue, total plant protein content, fruit protein content, seed protein content, protein content in a vegetative tissue, drought tolerance, nitrogen uptake, root lodging, stalk lodging, root penetration, plant height, ear length, and harvest index.

[0129] "Harvest index" refers to the grain weight divided by the total plant weight.

[0130] "rt1" (rootless 1) refers to the nucleotide sequence of the Zea Mays mutant.

[0131] "rt1" refers to the polypeptide of the Zea Mays mutant.

[0132] "RT1" refers to the Zea Mays RT1 wild type gene and cDNA and includes without limitation SEQ ID NO:9 and SEQ ID NO:12, respectively. "RT1" refers to the Zea Mays RT1 wild type protein encoded by the exons of SEQ ID NO:9 and by the cDNA of SEQ ID NO:12.

[0133] "RT1-like" refers to the nucleotide homologs of the maize RT1 sequence and corresponds to a rice, Arabidopsis, and additional maize sequence including without limitation the nucleotide sequences of SEQ ID NO:16, 18, and 20, respectively.

[0134] "RT1-like" refers to the polypeptide homologs of the maize RT1 protein and include without limitation the amino acid sequences of SEQ ID NO:17, 19 and 21, corre-sponding to an additional rice, Arabidopsis and additional maize homolog, respectively.

[0135] "Environmental conditions" refer to conditions under which the plant is grown, such as the availability of water, availability of nutrients (for example nitrogen or phos-phate), the soil type, or the presence of insects or disease.

[0136] "Varying environmental conditions" refer to changes in the environmental conditions under which the plant is grown, including and not limited to water availability, nutrient availability (for example nitrogen or phosphate), soil type, or presence of insects or disease.

[0137] "Root lodging" refers to stalks leaning from the center. Root lodging can occur as early as the late vegetative stages and as late as harvest maturity. Root lodging can be affected by hybrid susceptibility (i.e. disposition of a hybrid to be affected by pests that result in root lodging), environ-mental stress (drought, flooding), insect and disease injury. Root lodging can be attributed to corn rootworm injury in some cases.

[0138] "Root penetration" refers to the rate and depth of penetration of the plant root into the soil.

[0139] "Soil type" refers in terms of soil texture to the different sizes of particles, including and not limited to min-eral particles, in a particular sample. The term "soil type" also refers to the compactness of the soil under changing physical conditions including and not limited to water content and tilling. In general, soil is made up in part of finely ground rock particles, grouped according to size as sand, silt, and clay. Each size plays a significantly different role. For example, the largest particles, sand, determine aeration and drainage char-acteristics, while the tiniest, sub-microscopic clay particles, are chemically active, binding with water and plant nutrients. The ratio of these sizes determines soil type: clay, loam, clay-loam, silt-loam, and so on.

[0140] "Genome" as it applies to plant cells encompasses not only chromosomal DNA found within the nucleus, but organelle DNA found within subcellular components (e.g., mitochondrial, plastid) of the cell.

[0141] "Plant" includes reference to whole plants, plant organs, plant tissues, seeds and plant cells and progeny of same. Plant cells include, without limitation, cells from seeds, suspension cultures, embryos, meristematic regions, callus tissue, leaves, roots, shoots, gametophytes, sporophytes, pol-len, and microspores.

[0142] "Progeny" comprises any subsequent generation of a plant.

[0143] "Transgenic" refers to any cell, cell line, callus, tissue, plant part or plant, the genome of which has been altered by the presence of a heterologous nucleic acid, such as a recombinant DNA construct, including those initial trans-genic events as well as those created by sexual crosses or

asexual propagation from the initial transgenic event. The term "transgenic" as used herein does not encompass the alteration of the genome (chromosomal or extra-chromosomal) by conventional plant breeding methods or by naturally occurring events such as random cross-fertilization, non-recombinant viral infection, non-recombinant bacterial transformation, non-recombinant transposition, or spontaneous mutation.

[0144] "Transgenic plant" includes reference to a plant which comprises within its genome a heterologous polynucleotide. Preferably, the heterologous polynucleotide is stably integrated within the genome such that the polynucleotide is passed on to successive generations. The heterologous polynucleotide may be integrated into the genome alone or as part of a recombinant DNA construct.

[0145] "Heterologous" with respect to sequence means a sequence that originates from a foreign species, or, if from the same species, is substantially modified from its native form in composition and/or genomic locus by deliberate human intervention.

[0146] "Polynucleotide", "nucleic acid sequence", "nucleotide sequence". and "nucleic acid fragment" are used interchangeably to refer to a polymer of RNA or DNA that is single- or double-stranded, optionally containing synthetic, non-natural or altered nucleotide bases. Nucleotides (usually found in their 5'-monophosphate form) are referred to by their single letter designation as follows: "A" for adenylate or deoxyadenylate (for RNA or DNA, respectively), "C" for cytidylate or deoxycytidylate, "G" for guanylate or deoxyguanylate, "U" for uridylate, "T" for deoxythymidylate, "R" for purines (A or G), "Y" for pyrimidines (C or T), "K" for G or T, "H" for A or C or T, "I" for inosine, and "N" for any nucleotide.

[0147] "Polypeptide", "peptide", "amino acid sequence" and "protein" are used interchangeably herein to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue is an artificial chemical analogue of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers. The terms "polypeptide", "peptide", "amino acid sequence" and "protein" are also inclusive of modifications including and not limited to, glycosylation, lipid attachment, sulfation, gamma-carboxylation of glutamic acid residues, hydroxylation and ADP-ribosylation.

[0148] "Messenger RNA (mRNA)" refers to the RNA that is without introns and that can be translated into protein by the cell.

[0149] "cDNA" refers to a DNA that is complementary to and synthesized from a mRNA template using the enzyme reverse transcriptase. The cDNA can be single-stranded or converted into the double-stranded form using the Klenow fragment of DNA polymerase I.

[0150] "Mature" protein refers to a post-translationally processed polypeptide; i.e., one from which any pre- or pro-peptides present in the primary translation product have been removed.

[0151] "Precursor" protein refers to the primary product of translation of mRNA; i.e., with pre- and pro-peptides still present. Pre- and pro-peptides may be and are not limited to intracellular localization signals.

[0152] "Isolated" refers to materials, such as nucleic acid molecules and/or proteins, which are substantially free or otherwise removed from components that normally accompany or interact with the materials in a naturally occurring environment. Isolated polynucleotides may be purified from a host cell in which they naturally occur. Conventional nucleic acid purification methods known to skilled artisans may be used to obtain isolated polynucleotides. The term also embraces recombinant polynucleotides and chemically synthesized polynucleotides.

[0153] "Recombinant" refers to an artificial combination of two otherwise separated segments of sequence, e.g., by chemical synthesis or by the manipulation of isolated segments of nucleic acids by genetic engineering techniques. "Recombinant" also includes reference to a cell or vector, that has been modified by the introduction of a heterologous nucleic acid and a cell derived from a cell so modified. It does not encompass the alteration of the cell or vector by naturally occurring events (e.g., spontaneous mutation, natural transformation/transduction/transposition) such as those occurring without deliberate human intervention.

[0154] "Recombinant DNA construct" refers to a combination of nucleic acid fragments that are not normally found together in nature. Accordingly, a recombinant DNA construct may comprise regulatory sequences and coding sequences that are derived from different sources, or regulatory sequences and coding sequences derived from the same source, but arranged in a manner different than that normally found in nature.

[0155] The terms "regulatory sequence(s)" and "regulatory element(s)" are used interchangeably herein.

[0156] "Regulatory sequences" refer to nucleotide sequences located upstream (5' non-coding sequences), within, or downstream (3' non-coding sequences) of a coding sequence, and which influence the transcription, RNA processing or stability, or translation of the associated coding sequence. Regulatory sequences may include and are not limited to, promoters, translation leader sequences, introns, polyadenylation recognition sequences and the like.

[0157] "Promoter" refers to a nucleic acid fragment capable of controlling transcription of another nucleic acid fragment.

[0158] "Promoter functional in a plant" is a promoter capable of controlling transcription in plant cells whether or not its origin is from a plant cell.

[0159] "Tissue-specific promoter" and "tissue-preferred promoter" are used interchangeably, and refer to a promoter that is expressed predominantly but not necessarily exclusively in one tissue or organ and that may also be expressed in one specific cell.

[0160] "Developmentally regulated promoter" refers to a promoter whose activity is determined by developmental events.

[0161] "Operably linked" refers to the association of nucleic acid fragments in a single fragment so that the function of one is regulated by the other. For example, a promoter is operably linked with a nucleic acid fragment when it is capable of regulating the transcription of that nucleic acid fragment.

[0162] "Expression" refers to the production of a functional product. For example, expression of a nucleic acid fragment may refer to transcription of the nucleic acid fragment (e.g., transcription resulting in mRNA or functional RNA) and/or translation of mRNA into a precursor or mature protein.

[0163] "Phenotype" means the detectable characteristics of a cell or organism.

[0164] "Introduced" in the context of inserting a nucleic acid fragment (e.g., a recombinant DNA construct) into a cell,

means "transfection" or "transformation" or "transduction" and includes reference to the incorporation of a nucleic acid fragment into a eukaryotic or prokaryotic cell where the nucleic acid fragment may be incorporated into the genome of the cell (e.g., chromosome, plasmid, plastid or mitochondrial DNA), converted into an autonomous replicon, or transiently expressed (e.g., transfected mRNA).

[0165] A "transformed cell" is any cell into which a nucleic acid fragment (e.g., a recombinant DNA construct) has been introduced.

[0166] "Transformation" as used herein refers to both stable transformation and transient transformation.

[0167] "Stable transformation" refers to the introduction of a nucleic acid fragment into a genome of a host organism resulting in genetically stable inheritance. Once stably transformed, the nucleic acid fragment is stably integrated in the genome of the host organism and any subsequent generation.

[0168] "Transient transformation" refers to the introduction of a nucleic acid fragment into the nucleus, or DNA-containing organelle, of a host organism resulting in gene expression without genetically stable inheritance.

[0169] "Allele" is one of several alternative forms of a gene occupying a given locus on a chromosome. When the alleles present at a given locus on a pair of homologous chromosomes in a diploid plant are the same that plant is homozygous at that locus. If the alleles present at a given locus on a pair of homologous chromosomes in a diploid plant differ that plant is heterozygous at that locus. If a transgene is present on one of a pair of homologous chromosomes in a diploid plant that plant is hemizygous at that locus.

[0170] Sequence alignments and percent identity calculations may be determined using a variety of comparison methods designed to detect homologous sequences including and not limited to, the Megalign® program of the LASAR-GENE® bioinformatics computing suite (DNASTAR® Inc., Madison, Wis.). Unless stated otherwise, multiple alignment of the sequences provided herein were performed using the Clustal V method of alignment (Higgins and Sharp (1989) *CABIOS.* 5:151-153) with the default parameters (GAP PENALTY=10, GAP LENGTH PENALTY=10). Default parameters for pairwise alignments and calculation of percent identity of protein sequences using the Clustal V method are KTUPLE=1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5. For nucleic acids these parameters are KTUPLE=2, GAP PENALTY=5, WINDOW=4 and DIAGONALS SAVED=4. After alignment of the sequences, using the Clustal V program, it is possible to obtain "percent identity" and "divergence" values by viewing the "sequence distances" table on the same program; unless stated otherwise, percent identities and divergences provided and claimed herein were calculated in this manner.

[0171] Standard recombinant DNA and molecular cloning techniques used herein are well known in the art and are described more fully in Sambrook, J., Fritsch, E. F. and Maniatis, T. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, 1989 (hereinafter "Sambrook").

[0172] Turning now to preferred embodiments:

[0173] Preferred embodiments include isolated polynucleotides and polypeptides, recombinant DNA constructs, compositions (such as plants or seeds) comprising these recombinant DNA constructs, and methods utilizing these recombinant DNA constructs.

[0174] Preferred Isolated Polynucleotides and Polypeptides

[0175] The present invention includes the following preferred isolated polynucleotides and polypeptides:

[0176] An isolated polynucleotide comprising: (i) a nucleic acid sequence encoding a polypeptide having an amino acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13 or 21 and wherein expression of said polypeptide in a plant results in an altered root architecture when compared to a control plant not comprising said recombinant DNA construct, or (ii) a full complement of the nucleic acid sequence of (i), wherein the full complement and the nucleic acid sequence of (i) consist of the same number of nucleotides and are 100% complementary.

[0177] Any of the foregoing isolated polynucleotides may be utilized in any recombinant DNA constructs (including suppression DNA constructs) of the present invention.

[0178] An isolated polypeptide having an amino acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13 or 21 and wherein expression of said polypeptide in a plant results in an altered plant root architecture when compared to a control plant not comprising said recombinant DNA construct.

[0179] An isolated polynucleotide comprising (i) a nucleic acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13 or 21 and wherein said polynucleotide encodes a polypeptide wherein expression of said polypeptide results in an altered root architecture when compared to a control plant not comprising said recombinant DNA construct or (ii) a full complement of the nucleic acid sequence of (i). Any of the foregoing isolated polynucleotides may be utilized in any recombinant DNA constructs (including suppression DNA constructs) of the present invention. The isolated polynucleotide encodes a RT1 or RT1-like protein.

[0180] Preferred Recombinant DNA Constructs and Suppression DNA Constructs

[0181] In one aspect, the present invention includes recombinant DNA constructs (including suppression DNA constructs).

[0182] In one preferred embodiment, a recombinant DNA construct comprises a polynucleotide operably linked to at least one regulatory sequence (e.g., a promoter functional in a plant), wherein the polynucleotide comprises (i) a nucleic acid sequence encoding an amino acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%,

70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13, 17, 19, or 21 or (ii) a full complement of the nucleic acid sequence of (i).

[0183] In another preferred embodiment, a recombinant DNA construct comprises a polynucleotide operably linked to at least one regulatory sequence (e.g., a promoter functional in a plant), wherein said polynucleotide comprises (i) a nucleic acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO:12, 16, 18, or 20 or (ii) a full complement of the nucleic acid sequence of (i).

[0184] FIGS. 3A-3B shows the multiple alignment of the full length amino acid sequences of B73 RT1 (SEQ ID NO:13), rice RT1 homolog (SEQ ID NO:17), *Arabiopsis* RT1 homolog (SEQ ID NO:19), and the maize RT1 homolog from clone cfp7n.pk6.13 (SEQ ID NO:21). Amino acids conserved among all sequences are indicated with an asterisk (*) on the top row; dashes are used by the program to maximize alignment of the sequences. Two highly conserved sequence motifs are shown underlined in the alignment. The method parameters used to produce the multiple alignment of the sequences below was performed using the Clustal method of alignment (Higgins and Sharp (1989) *CABIOS.* 5:151-153) with the default parameters (GAP PENALTY=10, GAP LENGTH PENALTY=10), and the pairwise alignment default parameters of KTUPLE=1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5.

[0185] FIG. 4 shows a chart of the percent sequence identity for each pair of amino acid sequences displayed in FIGS. 3A-3B.

[0186] In another preferred embodiment, a recombinant DNA construct comprises a polynucleotide operably linked to at least one regulatory sequence (e.g., a promoter functional in a plant), wherein said polynucleotide encodes a RT1 or RT1-like protein. Preferably, the RT1 or RT1-like protein is from *Arabidopsis thaliana, Zea mays, Glycine max, Glycine tabacina, Glycine soja* and *Glycine tomentella.*

[0187] In another aspect, the present invention includes suppression DNA constructs.

[0188] A suppression DNA construct preferably comprises at least one regulatory sequence (preferably a promoter functional in a plant) operably linked to (a) all or part of (i) a nucleic acid sequence encoding a polypeptide having an amino acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13 or 21 or (ii) a full complement of the nucleic acid sequence of (a)(i); or (b) a region derived from all or part of a sense strand or antisense strand of a target gene of interest, said region having a nucleic acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%,

70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to said all or part of a sense strand or antisense strand from which said region is derived, and wherein said target gene of interest encodes a RT1 protein; or (c) all or part of (i) a nucleic acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO:12 or 20 or (ii) a full complement of the nucleic acid sequence of (c)(i). The suppression DNA construct preferably comprises a cosuppression construct, antisense construct, viral-suppression construct, hairpin suppression construct, stem-loop suppression construct, double-stranded RNA-producing construct, RNAi construct, or small RNA construct (e.g., an siRNA construct or an miRNA construct).

[0189] It is understood, as those skilled in the art will appreciate, that the invention encompasses more than the specific exemplary sequences. Alterations in a nucleic acid fragment which result in the production of a chemically equivalent amino acid at a given site, but do not affect the functional properties of the encoded polypeptide, are well known in the art. For example, a codon for the amino acid alanine, a hydrophobic amino acid, may be substituted by a codon encoding another less hydrophobic residue, such as glycine, or a more hydrophobic residue, such as valine, leucine, or isoleucine. Similarly, changes which result in substitution of one negatively charged residue for another, such as aspartic acid for glutamic acid, or one positively charged residue for another, such as lysine for arginine, can also be expected to produce a functionally equivalent product. Nucleotide changes which result in alteration of the N-terminal and C-terminal portions of the polypeptide molecule would also not be expected to alter the activity of the polypeptide. Each of the proposed modifications is well within the routine skill in the art, as is determination of retention of biological activity of the encoded products.

[0190] "Suppression DNA construct" is a recombinant DNA construct which when transformed or stably integrated into the genome of the plant, results in "silencing" of a target gene in the plant. The target gene may be endogenous or transgenic to the plant. "Silencing," as used herein with respect to the target gene, refers generally to the suppression of levels of mRNA or protein/enzyme expressed by the target gene, and/or the level of the enzyme activity or protein functionality. The terms "suppression", "suppressing" and "silencing", used interchangeably herein, include lowering, reducing, declining, decreasing, inhibiting, eliminating or preventing. "Silencing" or "gene silencing" does not specify mechanism and is inclusive, and not limited to, anti-sense, cosuppression, viral-suppression, hairpin suppression, stem-loop suppression, RNAi-based approaches, and small RNA-based approaches.

[0191] A suppression DNA construct may comprise a region derived from a target gene of interest and may comprise all or part of the nucleic acid sequence of the sense strand (or antisense strand) of the target gene of interest. Depending upon the approach to be utilized, the region may

be 100% identical or less than 100% identical (e.g., at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% identical) to all or part of the sense strand (or antisense strand) of the gene of interest.

[0192] Suppression DNA constructs are well-known in the art, are readily constructed once the target gene of interest is selected, and include, without limitation, cosuppression constructs, antisense constructs, viral-suppression constructs, hairpin suppression constructs, stem-loop suppression constructs, double-stranded RNA-producing constructs, and more generally, RNAi (RNA interference) constructs and small RNA constructs such as siRNA (short interfering RNA) constructs and miRNA (microRNA) constructs.

[0193] "Antisense inhibition" refers to the production of antisense RNA transcripts capable of suppressing the expression of the target protein.

[0194] "Antisense RNA" refers to an RNA transcript that is complementary to all or part of a target primary transcript or mRNA and that blocks the expression of a target isolated nucleic acid fragment (U.S. Pat. No. 5,107,065). The complementarity of an antisense RNA may be with any part of the specific gene transcript, i.e., at the 5' non-coding sequence, 3' non-coding sequence, introns, or the coding sequence.

[0195] "Cosuppression" refers to the production of sense RNA transcripts capable of suppressing the expression of the target protein. "Sense" RNA refers to RNA transcript that includes the mRNA and can be translated into protein within a cell or in vitro. Cosuppression constructs in plants have been previously designed by focusing on overexpression of a nucleic acid sequence having homology to a native mRNA, in the sense orientation, which results in the reduction of all RNA having homology to the overexpressed sequence (see Vaucheret et al. (1998) *Plant J.* 16:651-659; and Gura (2000) *Nature* 404:804-808).

[0196] Another variation describes the use of plant viral sequences to direct the suppression of proximal mRNA encoding sequences (PCT Publication WO 98/36083 published on Aug. 20, 1998).

[0197] Previously described is the use of "hairpin" structures that incorporate all, or part, of an mRNA encoding sequence in a complementary orientation that results in a potential "stem-loop" structure for the expressed RNA (PCT Publication WO 99/53050 published on Oct. 21, 1999). In this case the stem is formed by polynucleotides corresponding to the gene of interest inserted in either sense or anti-sense orientation with respect to the promoter and the loop is formed by some polynucleotides of the gene of interest, which do not have a complement in the construct. This increases the frequency of cosuppression or silencing in the recovered transgenic plants. For review of hairpin suppression see Wesley, S. V. et al. (2003) Methods in Molecular Biology, Plant Functional Genomics: Methods and Protocols 236:273-286.

[0198] A construct where the stem is formed by at least 30 nucleotides from a gene to be suppressed and the loop is formed by a random nucleotide sequence has also effectively been used for suppression (PCT Publication No. WO 99/61632 published on Dec. 2, 1999).

[0199] The use of poly-T and poly-A sequences to generate the stem in the stem-loop structure has also been described (PCT Publication No. WO 02/00894 published Jan. 3, 2002).

[0200] Yet another variation includes using synthetic repeats to promote formation of a stem in the stem-loop structure. Transgenic organisms prepared with such recombinant DNA fragments have been shown to have reduced levels of the protein encoded by the nucleotide fragment forming the loop as described in PCT Publication No. WO 02/00904, published 3 Jan. 2002.

[0201] RNA interference refers to the process of sequence-specific post-transcriptional gene silencing in animals mediated by short interfering RNAs (siRNAs) (Fire et al., Nature 391:806 1998). The corresponding process in plants is commonly referred to as post-transcriptional gene silencing (PTGS) or RNA silencing and is also referred to as quelling in fungi. The process of post-transcriptional gene silencing is thought to be an evolutionarily-conserved cellular defense mechanism used to prevent the expression of foreign genes and is commonly shared by diverse flora and phyla (Fire et al., Trends Genet. 15:358 1999). Such protection from foreign gene expression may have evolved in response to the production of double-stranded RNAs (dsRNAs) derived from viral infection or from the random integration of transposon elements into a host genome via a cellular response that specifically destroys homologous single-stranded RNA of viral genomic RNA. The presence of dsRNA in cells triggers the RNAi response through a mechanism that has yet to be fully characterized.

[0202] The presence of long dsRNAs in cells stimulates the activity of a ribonuclease III enzyme referred to as dicer. Dicer is involved in the processing of the dsRNA into short pieces of dsRNA known as short interfering RNAs (siRNAs) (Berstein et al., Nature 409:363 2001). Short interfering RNAs derived from dicer activity are typically about 21 to about 23 nucleotides in length and comprise about 19 base pair duplexes (Elbashir et al., Genes Dev. 15:188 2001). Dicer has also been implicated in the excision of 21- and 22-nucleotide small temporal RNAs (stRNAs) from precursor RNA of conserved structure that are implicated in translational control (Hutvagner et al., 2001, Science 293:834). The RNAi response also features an endonuclease complex, commonly referred to as an RNA-induced silencing complex (RISC), which mediates cleavage of single-stranded RNA having sequence complementarity to the antisense strand of the siRNA duplex. Cleavage of the target RNA takes place in the middle of the region complementary to the antisense strand of the siRNA duplex (Elbashir et al., Genes Dev. 15:188 2001). In addition, RNA interference can also involve small RNA (e.g., miRNA) mediated gene silencing, presumably through cellular mechanisms that regulate chromatin structure and thereby prevent transcription of target gene sequences (see, e.g., Allshire, Science 297:1818-1819 2002; Volpe et al., Science 297:1833-1837 2002; Jenuwein, Science 297:2215-2218 2002; and Hall et al., Science 297:2232-2237 2002). As such, miRNA molecules of the invention can be used to mediate gene silencing via interaction with RNA transcripts or alternately by interaction with particular gene sequences, wherein such interaction results in gene silencing either at the transcriptional or post-transcriptional level.

[0203] RNAi has been studied in a variety of systems. Fire et al. (Nature 391:806 1998) were the first to observe RNAi in *C. elegans*. Wianny and Goetz (Nature Cell Biol. 2:70 1999) describe RNAi mediated by dsRNA in mouse embryos. Ham-

mond et al. (Nature 404:293 2000) describe RNAi in *Drosophila* cells transfected with dsRNA. Elbashir et al., (Nature 411:494 2001) describe RNAi induced by introduction of duplexes of synthetic 21-nucleotide RNAs in cultured mammalian cells including human embryonic kidney and HeLa cells.

[0204] Small RNAs play an important role in controlling gene expression. Regulation of many developmental processes, including flowering, is controlled by small RNAs. It is now possible to engineer changes in gene expression of plant genes by using transgenic constructs which produce small RNAs in the plant.

[0205] Small RNAs appear to function by base-pairing to complementary RNA or DNA target sequences. When bound to RNA, small RNAs trigger either RNA cleavage or translational inhibition of the target sequence. When bound to DNA target sequences, it is thought that small RNAs can mediate DNA methylation of the target sequence. The consequence of these events, regardless of the specific mechanism, is that gene expression is inhibited.

[0206] It is thought that sequence complementarity between small RNAs and their RNA targets helps to determine which mechanism, RNA cleavage or translational inhibition, is employed. It is believed that siRNAs, which are perfectly complementary with their targets, work by RNA cleavage. Some miRNAs have perfect or near-perfect complementarity with their targets, and RNA cleavage has been demonstrated for at least a few of these miRNAs. Other miRNAs have several mismatches with their targets, and apparently inhibit their targets at the translational level. Again, without being held to a particular theory on the mechanism of action, a general rule is emerging that perfect or near-perfect complementarity causes RNA cleavage, whereas translational inhibition is favored when the miRNA/target duplex contains many mismatches. The apparent exception to this is microRNA 172 (miR172) in plants. One of the targets of miR172 is APETALA2 (AP2), and although miR172 shares near-perfect complementarity with AP2 it appears to cause translational inhibition of AP2 rather than RNA cleavage.

[0207] MicroRNAs (miRNAs) are noncoding RNAs of about 19 to about 24 nucleotides (nt) in length that have been identified in both animals and plants (Lagos-Quintana et al., Science 294:853-858 2001, Lagos-Quintana et al., Curr. Biol. 12:735-739 2002; Lau et al., Science 294:858-862 2001; Lee and Ambros, Science 294:862-864 2001; Llave et al., Plant Cell 14:1605-1619 2002; Mourelatos et al., Genes. Dev. 16:720-728 2002; Park et al., Curr. Biol. 12:1484-1495 2002; Reinhart et al., Genes. Dev. 16:1616-1626 2002). They are processed from longer precursor transcripts that range in size from approximately 70 to 200 nt, and these precursor transcripts have the ability to form stable hairpin structures. In animals, the enzyme involved in processing miRNA precursors is called Dicer, an RNAse III-like protein (Grishok et al., Cell 106:23-34 2001; Hutvagner et al., Science 293:834-838 2001; Ketting et al., Genes. Dev. 15:2654-2659 2001). Plants also have a Dicer-like enzyme, DCL1 (previously named CARPEL FACTORY/SHORT INTEGUMENTS1/SUSPENSOR1), and recent evidence indicates that it, like Dicer, is involved in processing the hairpin precursors to generate mature miRNAs (Park et al., Curr. Biol. 12:1484-1495 2002; Reinhart et al., Genes. Dev. 16:1616-1626 2002). Furthermore, it is becoming clear from recent work that at least some miRNA hairpin precursors originate as longer polyadeny-

lated transcripts, and several different miRNAs and associated hairpins can be present in a single transcript (Lagos-Quintana et al., Science 294:853-858 2001; Lee et al., EMBO J 21:4663-4670 2002). Recent work has also examined the selection of the miRNA strand from the dsRNA product arising from processing of the hairpin by DICER (Schwartz, et al. 2003 Cell 115:199-208). It appears that the stability (i.e. G:C vs. A:U content, and/or mismatches) of the two ends of the processed dsRNA affects the strand selection, with the low stability end being easier to unwind by a helicase activity. The 5' end strand at the low stability end is incorporated into the RISC complex, while the other strand is degraded.

[0208] MicroRNAs appear to regulate target genes by binding to complementary sequences located in the transcripts produced by these genes. In the case of lin-4 and let-7, the target sites are located in the 3' UTRs of the target mRNAs (Lee et al., Cell 75:843-854 1993; Wightman et al., Cell 75:855-862 1993; Reinhart et al., Nature 403:901-906 2000; Slack et al., Mol. Cell 5:659-669 2000), and there are several mismatches between the lin-4 and let-7 miRNAs and their target sites. Binding of the lin-4 or let-7 miRNA appears to cause downregulation of steady-state levels of the protein encoded by the target mRNA without affecting the transcript itself (Olsen and Ambros, Dev. Biol. 216:671-680 1999). On the other hand, recent evidence suggests that miRNAs can in some cases cause specific RNA cleavage of the target transcript within the target site, and this cleavage step appears to require 100% complementarity between the miRNA and the target transcript (Hutvagner and Zamore, Science 297:2056-2060 2002; Llave et al., Plant Cell 14:1605-1619 2002). It seems likely that miRNAs can enter at least two pathways of target gene regulation: Protein downregulation when target complementarity is <100%, and RNA cleavage when target complementarity is 100%. MicroRNAs entering the RNA cleavage pathway are analogous to the 21-25 nt short interfering RNAs (siRNAs) generated during RNA interference (RNAi) in animals and posttranscriptional gene silencing (PTGS) in plants (Hamilton and Baulcombe 1999; Hammond et al., 2000; Zamore et al., 2000; Elbashir et al., 2001), and likely are incorporated into an RNA-induced silencing complex (RISC) that is similar or identical to that seen for RNAi.

[0209] Identifying the targets of miRNAs with bioinformatics has not been successful in animals, and this is probably due to the fact that animal miRNAs have a low degree of complementarity with their targets. On the other hand, bioinformatic approaches have been successfully used to predict targets for plant miRNAs (Llave et al., Plant Cell 14:1605-1619 2002; Park et al., Curr. Biol. 12:1484-1495 2002; Rhoades et al., Cell 110:513-520 2002), and thus it appears that plant miRNAs have higher overall complementarity with their putative targets than do animal miRNAs. Most of these predicted target transcripts of plant miRNAs encode members of transcription factor families implicated in plant developmental patterning or cell differentiation.

[0210] A recombinant DNA construct (including a suppression DNA construct) of the present invention preferably comprises at least one regulatory sequence.

[0211] A preferred regulatory sequence is a promoter.

[0212] A number of promoters can be used in recombinant DNA constructs (and suppression DNA constructs) of the present invention. The promoters can be selected based on the desired outcome, and may include constitutive, tissue-specific, inducible, or other promoters for expression in the host organism.

11

[0213] High level, constitutive expression of the candidate gene under control of the 35S promoter may have pleiotropic effects. Candidate gene efficacy may be tested when driven by different promoters.

[0214] Suitable constitutive promoters for use in a plant host cell include, for example, the core promoter of the Rsyn7 promoter and other constitutive promoters disclosed in WO 99/43838 and U.S. Pat. No. 6,072,050; the core CaMV 35S promoter (Odell et al., Nature 313:810-812 (1985)); rice actin (McElroy et al., Plant Cell 2:163-171 (1990)); ubiquitin (Christensen et al., Plant Mol. Biol. 12:619-632 (1989) and Christensen et al., Plant Mol. Biol. 18:675-689 (1992)); pEMU (Last et al., Theor. Appl. Genet. 81:581-588 (1991)); MAS (Velten et al., EMBO J. 3:2723-2730 (1984)); ALS promoter (U.S. Pat. No. 5,659,026), and the like. Other constitutive promoters include, for example, those discussed in U.S. Pat. Nos. 5,608,149; 5,608,144; 5,604,121; 5,569,597; 5,466,785; 5,399,680; 5,268,463; 5,608,142; and 6,177,611 and maize GOS2 (WO0020571 A2).

[0215] In choosing a promoter to use in the methods of the invention, it may be desirable to use a tissue-specific or developmentally regulated promoter.

[0216] A preferred tissue-specific or developmentally regulated promoter is a DNA sequence which regulates the expression of a DNA sequence selectively in the cells/tissues of a plant critical to tassel development, seed set, or both, and limits the expression of such a DNA sequence to the period of tassel development or seed maturation in the plant. Any identifiable promoter may be used in the methods of the present invention which causes the desired temporal and spatial expression.

[0217] Promoters which are seed or embryo specific and may be useful in the invention include soybean Kunitz trysin inhibitor (Kti3, Jofuku and Goldberg, Plant Cell 1:1079-1093 (1989)), patatin (potato tubers) (Rocha-Sosa, M., et al. (1989) EMBO J. 8:23-29), convicilin, vicilin, and legumin (pea cotyledons) (Rerie, W. G., et al. (1991) Mol. Gen. Genet. 259: 149-157; Newbigin, E. J., et al. (1990) Planta 180:461-470; Higgins, T. J. V., et al. (1988) Plant. Mol. Biol. 11:683-695), zein (maize endosperm) (Schemthaner, J. P., et al. (1988) EMBO J. 7:1249-1255), phaseolin (bean cotyledon) (Segupta-Gopalan, C., et al. (1985) Proc. Natl. Acad. Sci. U.S.A. 82:3320-3324), phytohemagglutinin (bean cotyledon) (Voelker, T. et al. (1987) EMBO J. 6:3571-3577), B-conglycinin and glycinin (soybean cotyledon) (Chen, Z-L, et al. (1988) EMBO J. 7:297-302), glutelin (rice endosperm), hordein (barley endosperm) (Marris, C., et al. (1988) Plant Mol. Biol. 10:359-366), glutenin and gliadin (wheat endosperm) (Colot, V., et al. (1987) EMBO J. 6:3559-3564), and sporamin (sweet potato tuberous root) (Hattori, T., et al. (1990) Plant Mol. Biol. 14:595-604). Promoters of seed-specific genes operably linked to heterologous coding regions in chimeric gene constructions maintain their temporal and spatial expression pattern in transgenic plants. Such examples include Arabidopsis thaliana 2S seed storage protein gene promoter to express enkephalin peptides in Arabidopsis and Brassica napus seeds (Vanderkerckhove et al., Bio/Technology 7:L929-932 (1989)), bean lectin and bean beta-phaseolin promoters to express luciferase (Riggs et al., Plant Sci. 63:47-57 (1989)), and wheat glutenin promoters to express chloramphenicol acetyl transferase (Colot et al., EMBO J 6:3559-3564 (1987)).

[0218] Inducible promoters selectively express an operably linked DNA sequence in response to the presence of an endogenous or exogenous stimulus, for example by chemical compounds (chemical inducers) or in response to environmental, hormonal, chemical, and/or developmental signals. Inducible or regulated promoters include, for example, promoters regulated by light, heat, stress, flooding or drought, phytohormones, wounding, or chemicals such as ethanol, jasmonate, salicylic acid, or safeners.

[0219] Preferred promoters include the following: 1) the stress-inducible RD29A promoter (Kasuga et al. (1999) Nature Biotechnol. 17:287-91); 2) the barley promoter, B22E; expression of B22E is specific to the pedicel in developing maize kernels ("Primary Structure of a Novel Barley Gene Differentially Expressed in Immature Aleurone Layers". Klemsdal, S. S. et al., Mol. Gen. Genet. 228(1/2):9-16 (1991)); and 3) maize promoter, Zag2 ("Identification and molecular characterization of ZAG1, the maize homolog of the Arabidopsis floral homeotic gene AGAMOUS", Schmidt, R. J. et al., Plant Cell 5(7):729-737 (1993)). "Structural characterization, chromosomal localization and phylogenetic evaluation of two pairs of AGAMOUS-like MADS-box genes from maize", Theissen et al., Gene 156(2): 155-166 (1995); NCBI GenBank Accession No. X80206)). Zag2 transcripts can be detected 5 days prior to pollination to 7 to 8 days after pollination (DAP), and directs expression in the carpel of developing female inflorescences and CimI which is specific to the nucleus of developing maize kernels. CimI transcript is detected 4 to 5 days before pollination to 6 to 8 DAP. Other useful promoters include any promoter which can be derived from a gene whose expression is maternally associated with developing female florets.

[0220] Additional preferred promoters for regulating the expression of the nucleotide sequences of the present invention in plants are stalk-specific promoters. Such stalk-specific promoters include the alfalfa S2A promoter (GenBank Accession No. EF030816; Abrahams et al., Plant Mol. Biol. 27:513-528 (1995)) and S2B promoter (GenBank Accession No. EF030817) and the like, herein incorporated by reference.

[0221] Promoters may be derived in their entirety from a native gene, or be composed of different elements derived from different promoters found in nature, or even comprise synthetic DNA segments. It is understood by those skilled in the art that different promoters may direct the expression of a gene in different tissues or cell types, or at different stages of development, or in response to different environmental conditions. It is further recognized that since in most cases the exact boundaries of regulatory sequences have not been completely defined, DNA fragments of some variation may have identical promoter activity. Promoters that cause a gene to be expressed in most cell types at most times are commonly referred to as "constitutive promoters". New promoters of various types useful in plant cells are constantly being discovered; numerous examples may be found in the compilation by Okamuro, J. K., and Goldberg, R. B., Biochemistry of Plants 15:1-82 (1989).

[0222] Preferred promoters may include: RIP2, mLIP15, ZmCOR1, Rab17, CaMV 35S, RD29A, B22E, Zag2, SAM synthetase, ubiquitin, CaMV 19S, nos, Adh, sucrose synthase, R-allele, root cell promoter, the vascular tissue preferred promoters S2A (Genbank accession number EF030816; SEQ ID NO:76) and S2B (Genbank accession number EF030817) and the constitutive promoter GOS2 from Zea mays. Other preferred promoters include root preferred promoters, such as the maize NAS2 promoter, the

maize Cyclo promoter (US 2006/0156439, published Jul. 13, 2006), the maize ROOTMET2 promoter (WO05063998, published Jul. 14, 2005), the CR1BIO promoter (WO06055487, published May 26, 2006), the CRWAQ81 (WO05035770, published Apr. 21, 2005) and the maize ZRP2.47 promoter (NCBI accession number: U38790, gi: 1063664),

[0223] Recombinant DNA constructs (and suppression DNA constructs) of the present invention may also include other regulatory sequences, including and not limited to, translation leader sequences, introns, and polyadenylation recognition sequences. In another preferred embodiment of the present invention, a recombinant DNA construct of the present invention further comprises an enhancer or silencer.

[0224] An intron sequence can be added to the 5' untranslated region or the coding sequence of the partial coding sequence to increase the amount of the mature message that accumulates in the cytosol. Inclusion of a spliceable intron in the transcription unit in both plant and animal expression constructs has been shown to increase gene expression at both the mRNA and protein levels up to 1000-fold. Buchman and Berg, *Mol. Cell Biol.* 8:4395-4405 (1988); Callis et al., *Genes Dev.* 1:1183-1200 (1987). Such intron enhancement of gene expression is typically greatest when placed near the 5' end of the transcription unit. Use of maize introns Adh1-S intron 1, 2, and 6, the Bronze-1 intron are known in the art. See generally, *The Maize Handbook*, Chapter 116, Freeling and Walbot, Eds., Springer, N.Y. (1994).

[0225] If polypeptide expression is desired, it is generally desirable to include a polyadenylation region at the 3'-end of a polynucleotide coding region. The polyadenylation region can be derived from the natural gene, from a variety of other plant genes, or from T-DNA. The 3' end sequence to be added can be derived from, for example, the nopaline synthase or octopine synthase genes, or alternatively from another plant gene, or less preferably from any other eukaryotic gene.

[0226] A translation leader sequence is a DNA sequence located between the promoter sequence of a gene and the coding sequence. The translation leader sequence is present in the fully processed mRNA upstream of the translation start sequence. The translation leader sequence may affect processing of the primary transcript to mRNA, mRNA stability or translation efficiency. Examples of translation leader sequences have been described (Turner, R. and Foster, G. D. Molecular Biotechnology 3:225 (1995)).

[0227] Any plant can be selected for the identification of regulatory sequences and genes to be used in creating recombinant DNA constructs and suppression DNA constructs of the present invention. Examples of suitable plant targets for the isolation of genes and regulatory sequences would include but are not limited to alfalfa, apple, apricot, *Arabidopsis*, artichoke, arugula, asparagus, avocado, banana, barley, beans, beet, blackberry, blueberry, broccoli, brussels sprouts, cabbage, canola, cantaloupe, carrot, cassaya, castorbean, cauliflower, celery, cherry, chicory, cilantro, citrus, clementines, clover, coconut, coffee, corn, cotton, cranberry, cucumber, Douglas fir, eggplant, endive, escarole, eucalyptus, fennel, figs, garlic, gourd, grape, grapefruit, honey dew, jicama, kiwifruit, lettuce, leeks, lemon, lime, Loblolly pine, linseed, mango, melon, millet, mushroom, nectarine, nut, oat, oil palm, oil seed rape, okra, olive, onion, orange, an ornamental plant, palm, papaya, parsley, parsnip, pea, peach, peanut, pear, pepper, persimmon, pine, pineapple, plantain, plum, pomegranate, poplar, potato, pumpkin, quince, radiata pine, radicchio, radish, rapeseed, raspberry, rice, rye, sorghum, Southern pine, soybean, spinach, squash, strawberry, sugarbeet, sugarcane, sunflower, sweet potato, sweetgum, tangerine, tea, tobacco, tomato, triticale, turf, turnip, a vine, watermelon, wheat, yams, and zucchini. Particularly preferred plants for the identification of regulatory sequences are *Arabidopsis*, corn, wheat, soybean, and cotton.

[0228] Preferred Compositions

[0229] A preferred composition of the present invention is a plant comprising in its genome any of the recombinant DNA constructs (including any of the suppression DNA constructs) of the present invention (such as those preferred constructs discussed above). A Preferred composition also includes any progeny of the plant, and any seed obtained from the plant or its progeny. Progeny includes subsequent generations obtained by self-pollination or out-crossing of a plant. Progeny also includes hybrids and inbreds.

[0230] Preferably, in hybrid seed propagated crops, mature transgenic plants can be self-pollinated to produce a homozygous inbred plant. The inbred plant produces seed containing the newly introduced recombinant DNA construct (or suppression DNA construct). These seeds can be grown to produce plants that would exhibit an altered agronomic characteristic (e.g. an increased agronomic characteristic under nitrogen or phosphate limiting conditions), or used in a breeding program to produce hybrid seed, which can be grown to produce plants that would exhibit altered root architecture. Preferably, the seeds are maize.

[0231] Preferably, the plant is a monocotyledonous or dicotyledonous plant, more preferably, a maize or soybean plant, even more preferably a maize plant, such as a maize hybrid plant or a maize inbred plant. The plant may also be sunflower, sorghum, canola, wheat, alfalfa, cotton, rice, barley or millet.

[0232] Preferably, the recombinant DNA construct is stably integrated into the genome of the plant.

[0233] Particularly preferred embodiments include but are not limited to the following preferred embodiments:

[0234] 1. A plant (preferably a maize or soybean plant) comprising in its genome a recombinant DNA construct comprising a polynucleotide operably linked to at least one regulatory sequence, wherein said polynucleotide encodes a polypeptide having an amino acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13, 17, 19, or 21, and wherein said plant exhibits an altered root architecture when compared to a control plant not comprising said recombinant DNA construct. Preferably, the plant further exhibits an alteration of at least one agronomic characteristic when compared to the control plant.

[0235] 2. A plant (preferably a maize or soybean plant) comprising in its genome a recombinant DNA construct comprising a polynucleotide operably linked to at least one regulatory sequence, wherein said polynucleotide encodes a RT1 or RT1-like protein, and wherein said plant exhibits an altered root architecture when compared to a control plant not comprising said recombinant DNA construct. Preferably, the plant further exhibits an alteration of at least one agronomic characteristic when compared to the control plant. Preferably,

the RT1 or RT1-like protein is from *Arabidopsis thaliana, Zea mays, Glycine max, Glycine tabacina, Glycine soja* or *Glycine tomentella.*

[0236] 3. A plant (preferably a maize or soybean plant) comprising in its genome a suppression DNA construct comprising at least one regulatory element operably linked to a region derived from all or part of a sense strand or antisense strand of a target gene of interest, said region having a nucleic acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to said all or part of a sense strand or antisense strand from which said region is derived, and wherein said target gene of interest encodes a RT1 or RT1-like protein, and wherein said plant exhibits an alteration of at least one agronomic characteristic when compared to a control plant not comprising said recombinant DNA construct.

[0237] 4. A plant (preferably a maize or soybean plant) comprising in its genome a suppression DNA construct comprising at least one regulatory element operably linked to all or part of (a) a nucleic acid sequence encoding a polypeptide having an amino acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13, 17, 19, or 21, or (b) a full complement of the nucleic acid sequence of (a), and wherein said plant exhibits an alteration of at least one agronomic characteristic when compared to a control plant not comprising said recombinant DNA construct.

[0238] 5. Any progeny of the above plants in preferred embodiments 1-4, any seeds of the above plants in preferred embodiments 1-4, any seeds of progeny of the above plants in preferred embodiments 1-4, and cells from any of the above plants in preferred embodiments 1-4 and progeny thereof.

[0239] In any of the foregoing preferred embodiments 1-5 or any other embodiments of the present invention, the recombinant DNA construct (or suppression DNA construct) preferably comprises at least a promoter that is functional in a plant as a preferred regulatory sequence.

[0240] In any of the foregoing preferred embodiments 1-5 or any other embodiments of the present invention, the alteration of at least one agronomic characteristic is either an increase or decrease, preferably an increase.

[0241] In any of the foregoing preferred embodiments 1-5 or any other embodiments of the present invention, the at least one greenness, yield, growth rate, biomass, fresh weight at maturation, dry weight at maturation, fruit yield, seed yield, total plant nitrogen content, fruit nitrogen content, seed nitrogen content, nitrogen content in a vegetative tissue, total plant free amino acid content, fruit free amino acid content, seed free amino acid content, free amino acid content in a vegetative tissue, total plant protein content, fruit protein content, seed protein content, protein content in a vegetative tissue, drought tolerance, nitrogen uptake, root lodging, root penetration and harvest index.

[0242] Greenness, harvest index, yield, biomass, resistance to root lodging, improved root penetration are particularly preferred agronomic characteristic for alteration. Further, these agronomic characteristics preferably are increased relative to the control.

[0243] In any of the foregoing preferred embodiments 1-5 or any other embodiments of the present invention, the plant preferably exhibits the alteration of at least one agronomic characteristic irrespective of the for example water and nutrient availability when compared to a control plant.

[0244] One of ordinary skill in the art is familiar with protocols for determining alteration in plant root architecture. For example, alterations in root architecture can be determined by counting the nodal root numbers of the top 3 or 4 nodes of the greenhouse grown plants or the width of the root band. Other measures of alterations in root architecture include but are not limited to alterations in vigor, growth, size, yield, biomass, improved root penetration or resistance to root lodging when compared to a control or reference plant.

[0245] The Examples below describe some representative protocols and techniques for detecting alterations in root architecture.

[0246] One can also evaluate alterations in root architecture by the ability of the plant to maintain sufficient yield thresholds in field testing under various environmental conditions (e.g. nutrient over-abundance or limitation, water over-abundance or limitation, exposure to insects or disease) by measuring for substantially equivalent yield at those conditions compared to normal nutrient or water conditions, or by measuring for less yield drag under over-abundant or limiting nutrient and water conditions compared to a control or reference plant.

[0247] Alterations in root architecture can also be measured by determining the resistance to root lodging of the transgenic plants compared to reference or control plant. Improved root penetration is an additional measure to determine alterations in root architecture.

[0248] One of ordinary skill in the art would readily recognize a suitable control or reference plant to be utilized when assessing or measuring an agronomic characteristic or phenotype of a transgenic plant in any embodiment of the present invention in which a control or reference plant is utilized (e.g., compositions or methods as described herein). For example, by way of non-limiting illustrations:

[0249] 1. Progeny of a transformed plant which is hemizygous with respect to a recombinant DNA construct (or suppression DNA construct), such that the progeny are segregating into plants either comprising or not comprising the recombinant DNA construct (or suppression DNA construct): the progeny comprising the recombinant DNA construct (or suppression DNA construct) would be typically measured relative to the progeny not comprising the recombinant DNA construct (or suppression DNA construct) (i.e., the progeny not comprising the recombinant DNA construct (or suppression DNA construct) is the control or reference plant).

[0250] 2. Introgression of a recombinant DNA construct (or suppression DNA construct) into an inbred line, such as in maize, or into a variety, such as in soybean: the introgressed line would typically be measured relative to the parent inbred or variety line (i.e., the parent inbred or variety line is the control or reference plant).

[0251] 3. Two hybrid lines, where the first hybrid line is produced from two parent inbred lines, and the second hybrid line is produced from the same two parent inbred lines except

that one of the parent inbred lines contains a recombinant DNA construct (or suppression DNA construct): the second hybrid line would typically be measured relative to the first hybrid line (i.e., the parent inbred or variety line is the control or reference plant).

[0252] 4. A plant comprising a recombinant DNA construct (or suppression DNA construct): the plant may be assessed or measured relative to a control plant not comprising the recombinant DNA construct (or suppression DNA construct) but otherwise having a comparable genetic background to the plant (e.g., sharing at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity of nuclear genetic material compared to the plant comprising the recombinant DNA construct (or suppression DNA construct)). There are many laboratory-based techniques available for the analysis, comparison and characterization of plant genetic backgrounds; among these are Isozyme Electrophoresis, Restriction Fragment Length Polymorphisms (RFLPs), Randomly Amplified Polymorphic DNAs (RAPDs), Arbitrarily Primed Polymerase Chain Reaction (AP-PCR), DNA Amplification Fingerprinting (DAF), Sequence Characterized Amplified Regions (SCARs), Amplified Fragment Length Polymorphisms (AFLP®s), and Simple Sequence Repeats (SSRs) which are also referred to as Microsatellites.

[0253] Furthermore, one of ordinary skill in the art would readily recognize that a suitable control or reference plant to be utilized when assessing or measuring an agronomic characteristic or phenotype of a transgenic plant would not include a plant that had been previously selected, via mutagenesis or transformation, for the desired agronomic characteristic or phenotype.

[0254] Preferred Methods

[0255] Preferred methods include but are not limited to methods for altering root architecture in a plant, methods for evaluating alteration of root architecture in a plant, methods for altering an agronomic characteristic in a plant, methods for evaluating an alteration of an agronomic characteristic in a plant, and methods for producing seed. Preferably, the plant is a monocotyledonous or dicotyledonous plant, more preferably, a maize or soybean plant, even more preferably a maize plant. The plant may also be sunflower, sorghum, canola, wheat, alfalfa, cotton, rice, barley or millet. The seed is preferably a maize or soybean seed, more preferably a maize seed, and even more preferably, a maize hybrid seed or maize inbred seed.

[0256] Particularly preferred methods include but are not limited to the following:

[0257] A method of altering root architecture of a plant, comprising: (a) introducing into a regenerable plant cell a recombinant DNA construct comprising a polynucleotide operably linked to at least one regulatory sequence (preferably a promoter functional in a plant), wherein the polynucleotide encodes a polypeptide having an amino acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13, 17, 19, or 21, and (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the recombinant DNA construct and

exhibits in altered root architecture when compared to a control plant not comprising the recombinant DNA construct. The method may further comprise (c) obtaining a progeny plant derived from the transgenic plant.

[0258] A method of altering root architecture in a plant, comprising: (a) introducing into a regenerable plant cell a suppression DNA construct comprising at least one regulatory sequence (preferably a promoter functional in a plant) operably linked to all or part of (i) a nucleic acid sequence encoding a polypeptide having an amino acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13, 17, 19, or 21 or (ii) a full complement of the nucleic acid sequence of (a)(i); and (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the recombinant DNA construct and exhibits an altered root architecture when compared to a control plant not comprising the recombinant DNA construct. The method may further comprise (c) obtaining a progeny plant derived from the transgenic plant.

[0259] A method of altering root architecture in a plant, comprising: (a) introducing into a regenerable plant cell a suppression DNA construct comprising at least one regulatory sequence (preferably a promoter functional in a plant) operably linked to a region derived from all or part of a sense strand or antisense strand of a target gene of interest, said region having a nucleic acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to said all or part of a sense strand or antisense strand from which said region is derived, and wherein said target gene of interest encodes a RT1 or RT1-like protein; and (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the recombinant DNA construct and exhibits an altered root architecture when compared to a control plant not comprising the recombinant DNA construct. The method may further comprise (c) obtaining a progeny plant derived from the transgenic plant.

[0260] A method of evaluating altered root architecture in a plant, comprising (a) introducing into a regenerable plant cell a recombinant DNA construct comprising a polynucleotide operably linked to at least on regulatory sequence (preferably a promoter functional in a plant), wherein the polynucleotide encodes a polypeptide having an amino acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13, 17, 19, or 21 (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the recombinant DNA construct; and

(c) evaluating the transgenic plant for altered root architecture compared to a control plant not comprising the recombinant DNA construct. The method may further comprise (d) obtaining a progeny plant derived from the transgenic plant, wherein the progeny plant comprises in its genome the recombinant DNA construct; and (e) evaluating the progeny plant for altered root architecture compared to a control plant not comprising the recombinant DNA construct.

[0261] A method of evaluating altered root architecture in a plant, comprising (a) introducing into a regenerable plant cell a suppression DNA construct comprising at least one regulatory sequence (preferably a promoter functional in a plant) operably linked to all or part of (i) a nucleic acid sequence encoding a polypeptide having an amino acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13, 17, 19, or 21, or (ii) a full complement of the nucleic acid sequence of (a)(i); (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the suppression DNA construct; and (c) evaluating the transgenic plant for altered root architecture compared to a control plant not comprising the suppression DNA construct. The method may further comprise (d) obtaining a progeny plant derived from the transgenic plant, wherein the progeny plant comprises in its genome the suppression DNA construct; and (e) evaluating the progeny plant for altered root architecture compared to a control plant not comprising the suppression DNA construct.

[0262] A method of evaluating altered root architecture in a plant, comprising (a) introducing into a regenerable plant cell a suppression DNA construct comprising at least one regulatory sequence (preferably a promoter functional in a plant) operably linked to a region derived from all or part of a sense strand or antisense strand of a target gene of interest, said region having a nucleic acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to said all or part of a sense strand or antisense strand from which said region is derived, and wherein said target gene of interest encodes a RT1 or RT1-like protein; (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the suppression DNA construct; and (c) evaluating the transgenic plant for altered root architecture compared to a control plant not comprising the suppression DNA construct. The method may further comprise (d) obtaining a progeny plant derived from the transgenic plant, wherein the progeny plant comprises in its genome the suppression DNA construct; and (e) evaluating the progeny plant for altered root architecture compared to a control plant not comprising the suppression DNA construct.

[0263] A method of evaluating altered root architecture in a plant, comprising (a) introducing into a regenerable plant cell a recombinant DNA construct comprising a polynucleotide operably linked to at least one regulatory sequence (prefer-

ably a promoter functional in a plant), wherein said polynucleotide encodes a polypeptide having an amino acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13, 17, 19, or 21 (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the recombinant DNA construct; (c) obtaining a progeny plant derived from said transgenic plant, wherein the progeny plant comprises in its genome the recombinant DNA construct; and (d) evaluating the progeny plant for altered root architecture compared to a control plant not comprising the recombinant DNA construct.

[0264] A method of evaluating altered root architecture in a plant, comprising (a) introducing into a regenerable plant cell a suppression DNA construct comprising at least one regulatory sequence (preferably a promoter functional in a plant) operably linked to all or part of (i) a nucleic acid sequence encoding a polypeptide having an amino acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13, 17, 19, or 21, or (ii) a full complement of the nucleic acid sequence of (a)(i); (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the suppression DNA construct; (c) obtaining a progeny plant derived from said transgenic plant, wherein the progeny plant comprises in its genome the suppression DNA construct; and (e) evaluating the progeny plant for altered root architecture compared to a control plant not comprising the suppression DNA construct.

[0265] A method of evaluating altered root architecture in a plant, comprising (a) introducing into a regenerable plant cell a suppression DNA construct comprising at least one regulatory sequence (preferably a promoter functional in a plant) operably linked to a region derived from all or part of a sense strand or antisense strand of a target gene of interest, said region having a nucleic acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to said all or part of a sense strand or antisense strand from which said region is derived, and wherein said target gene of interest encodes a RT1 or RT1-like protein; (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the suppression DNA construct; (c) obtaining a progeny plant derived from the transgenic plant, wherein the progeny plant comprises in its genome the suppression DNA construct; and (d) evaluating the progeny plant for altered root architecture compared to a control plant not comprising the recombinant DNA construct.

[0266] A method of evaluating an alteration of an agronomic characteristic in a plant, comprising (a) introducing into a regenerable plant cell a recombinant DNA construct comprising a polynucleotide operably linked to at least on regulatory sequence (preferably a promoter functional in a plant), wherein said polynucleotide encodes a polypeptide having an amino acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13, 17, 19, or 21 (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome said recombinant DNA construct; and (c) determining whether the transgenic plant exhibits an alteration in at least one agronomic characteristic when compared to a control plant not comprising the recombinant DNA construct. The method may further comprise (d) obtaining a progeny plant derived from the transgenic plant, wherein the progeny plant comprises in its genome the recombinant DNA construct; and (e) determining whether the progeny plant exhibits an alteration in at least one agronomic characteristic when compared to a control plant not comprising the recombinant DNA construct.

[0267] A method of evaluating an alteration of an agronomic characteristic in a plant, comprising (a) introducing into a regenerable plant cell a suppression DNA construct comprising at least one regulatory sequence (preferably a promoter functional in a plant) operably linked to all or part of (i) a nucleic acid sequence encoding a polypeptide having an amino acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13, 16, 17, or 19, or (ii) a full complement of the nucleic acid sequence of (i); (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the suppression DNA construct; and (c) determining whether the transgenic plant exhibits an alteration in at least one agronomic characteristic when compared to a control plant not comprising the suppression DNA construct. The method may further comprise (d) obtaining a progeny plant derived from the transgenic plant, wherein the progeny plant comprises in its genome the suppression DNA construct; and (e) determining whether the progeny plant exhibits an alteration in at least one agronomic characteristic when compared to a control plant not comprising the suppression DNA construct.

[0268] A method of evaluating alteration of an agronomic characteristic in a plant, comprising (a) introducing into a regenerable plant cell a suppression DNA construct comprising at least one regulatory sequence (preferably a promoter functional in a plant) operably linked to a region derived from all or part of a sense strand or antisense strand of a target gene of interest, said region having a nucleic acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%,

89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to said all or part of a sense strand or antisense strand from which said region is derived, and wherein said target gene of interest encodes RT1 protein; (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the suppression DNA construct; and (c) determining whether the transgenic plant exhibits an alteration in at least one agronomic characteristic when compared to a control plant not comprising the suppression DNA construct. The method may further comprise (d) obtaining a progeny plant derived from the transgenic plant, wherein the progeny plant comprises in its genome the suppression DNA construct; and (e) determining whether the progeny plant exhibits an alteration in at least one agronomic characteristic when compared to a control plant not comprising the suppression DNA construct.

[0269] A method of evaluating an alteration of an agronomic characteristic in a plant, comprising (a) introducing into a regenerable plant cell a recombinant DNA construct comprising a polynucleotide operably linked to at least one regulatory sequence (preferably a promoter functional in a plant), wherein said polynucleotide encodes a polypeptide having an amino acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13, 16, 17, or 19 (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome said recombinant DNA construct; (c) obtaining a progeny plant derived from said transgenic plant, wherein the progeny plant comprises in its genome the recombinant DNA construct; and (d) determining whether the progeny plant exhibits an alteration in at least one agronomic characteristic when compared to a control plant not comprising the recombinant DNA construct.

[0270] A method of evaluating an alteration of an agronomic characteristic in a plant, comprising (a) introducing into a regenerable plant cell a suppression DNA construct comprising at least one regulatory sequence (preferably a promoter functional in a plant) operably linked to all or part of (i) a nucleic acid sequence encoding a polypeptide having an amino acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13, 17, 19, or 21 or (ii) a full complement of the nucleic acid sequence of (i); (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the suppression DNA construct; (c) obtaining a progeny plant derived from said transgenic plant, wherein the progeny plant comprises in its genome the suppression DNA construct; and (d) determining whether the progeny plant exhibits an alteration in at least one agronomic characteristic when compared to a control plant not comprising the recombinant DNA construct.

[0271] A method of evaluating an alteration of an agronomic characteristic in a plant, comprising (a) introducing into a regenerable plant cell a suppression DNA construct comprising at least one regulatory sequence (preferably a promoter functional in a plant) operably linked to a region derived from all or part of a sense strand or antisense strand of a target gene of interest, said region having a nucleic acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to said all or part of a sense strand or antisense strand from which said region is derived, and wherein said target gene of interest encodes a RT1 protein; (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the suppression DNA construct; (c) obtaining a progeny plant derived from said transgenic plant, wherein the progeny plant comprises in its genome the suppression DNA construct; and (d) determining whether the progeny plant exhibits an alteration in at least one agronomic characteristic when compared to a control plant not comprising the suppression DNA construct.

[0272] A method of producing seed (preferably seed that can be sold as a product offering with altered root architecture) comprising any of the preceding preferred methods, and further comprising obtaining seeds from said progeny plant, wherein said seeds comprise in their genome said recombinant DNA construct (or suppression DNA construct).

[0273] In any of the preceding preferred methods, in said introducing step said regenerable plant cell preferably comprises a callus cell (preferably embryogenic), a gametic cell, a meristematic cell, or a cell of an immature embryo. The regenerable plant cells are preferably from an inbred maize plant.

[0274] In any of the preceding preferred methods or any other embodiments of methods of the present invention, said regenerating step preferably comprises: (i) culturing said transformed plant cells in a media comprising an embryogenic promoting hormone until callus organization is observed; (ii) transferring said transformed plant cells of step (i) to a first media which includes a tissue organization promoting hormone; and (iii) subculturing said transformed plant cells after step (ii) onto a second media, to allow for shoot elongation, root development or both.

[0275] The introduction of recombinant DNA constructs of the present invention into plants may be carried out by any suitable technique, including and not limited to direct DNA uptake, chemical treatment, electroporation, microinjection, cell fusion, infection, vector mediated DNA transfer, bombardment, or *Agrobacterium* mediated transformation.

[0276] In any of the preceding preferred methods or any other embodiments of methods of the present invention, the at least one agronomic characteristic is preferably selected from the group consisting of greenness, yield, growth rate, biomass, fresh weight at maturation, dry weight at maturation, fruit yield, seed yield, total plant nitrogen content, fruit nitrogen content, seed nitrogen content, nitrogen content in a vegetative tissue, total plant free amino acid content, fruit free amino acid content, seed free amino acid content, free amino acid content in a vegetative tissue, total plant protein content, fruit protein content, seed protein content, protein content in a

vegetative tissue, drought tolerance, nitrogen uptake, root lodging, stalk lodging, plant height, ear length, and harvest index; with greenness, yield, biomass, improved root penetration or resistance to root lodging being a particularly preferred agronomic characteristic for alteration (preferably an increase).

[0277] In any of the preceding preferred methods or any other embodiments of methods of the present invention, the plant preferably exhibits the alteration of at least one agronomic characteristic irrespective of the environmental conditions when compared to a control plant (e.g., water, nutrient availability, insect or disease),

[0278] The introduction of recombinant DNA constructs of the present invention into plants may be carried out by any suitable technique, including and not limited to direct DNA uptake, chemical treatment, electroporation, microinjection, cell fusion, infection, vector mediated DNA transfer, bombardment, or *Agrobacterium* mediated transformation.

[0279] Preferred techniques are set forth below in the Examples.

[0280] Other preferred methods for transforming dicots, primarily by use of *Agrobacterium tumefaciens*, and obtaining transgenic plants include those published for cotton (U.S. Pat. No. 5,004,863, U.S. Pat. No. 5,159,135, U.S. Pat. No. 5,518,908); soybean (U.S. Pat. No. 5,569,834, U.S. Pat. No. 5,416,011, McCabe et. al., *Bio/Technology* 6:923 (1988), Christou et al., *Plant Physiol.* 87:671 674 (1988)); *Brassica* (U.S. Pat. No. 5,463,174); peanut (Cheng et al., *Plant Cell Rep.* 15:653 657 (1996), McKently et al., *Plant Cell Rep.* 14:699 703 (1995)); papaya; and pea (Grant et al., *Plant Cell Rep.* 15:254 258, (1995)).

[0281] Transformation of monocotyledons using electroporation, particle bombardment, and *Agrobacterium* have also been reported and are included as preferred methods, for example, transformation and plant regeneration as achieved in asparagus (Bytebier et al., *Proc. Natl. Acad. Sci. U.S.A.* 84:5354, (1987)); barley (Wan and Lemaux, *Plant Physiol.* 104:37 (1994)); *Zea mays* (Rhodes et al., *Science* 240:204 (1988), Gordon-Kamm et al., *Plant Cell* 2:603 618 (1990), Fromm et al., *Bio/Technology* 8:833 (1990), Koziel et al., *Bio/Technology* 11:194, (1993), Armstrong et al., *Crop Science* 35:550-557 (1995)); oat (Somers et al., *Bio/Technology* 10:1589 (1992)); orchard grass (Horn et al., *Plant Cell Rep.* 7:469 (1988)); rice (Toriyama et al., *Theor. Appl. Genet.* 205:34, (1986); Part et al., *Plant Mol. Biol.* 32:1135 1148, (1996); Abedinia et al., *Aust. J. Plant Physiol.* 24:133 141 (1997); Zhang and Wu, *Theor. Appl. Genet.* 76:835 (1988); Zhang et al., *Plant Cell Rep.* 7:379, (1988); Battraw and Hall, *Plant Sci.* 86:191 202 (1992); Christou et al., *Bio/Technology* 9:957 (1991)); rye (De la Pena et al., *Nature* 325:274 (1987)); sugarcane (Bower and Birch, Plant J. 2:409 (1992)); tall fescue (Wang et al., *Bio/Technology* 10:691 (1992)), and wheat (Vasil et al., *Bio/Technology* 10:667 (1992); U.S. Pat. No. 5,631,152).

[0282] There are a variety of methods for the regeneration of plants from plant tissue. The particular method of regeneration will depend on the starting plant tissue and the particular plant species to be regenerated.

[0283] The regeneration, development, and cultivation of plants from single plant protoplast transformants or from various transformed explants is well known in the art (Weissbach and Weissbach, In: Methods for Plant Molecular Biology, (Eds.), Academic Press, Inc. San Diego, Calif., (1988)). This regeneration and growth process typically includes the

steps of selection of transformed cells, culturing those individualized cells through the usual stages of embryonic development through the rooted plantlet stage. Transgenic embryos and seeds are similarly regenerated. The resulting transgenic rooted shoots are thereafter planted in an appropriate plant growth medium such as soil.

[0284] The development or regeneration of plants containing the foreign, exogenous isolated nucleic acid fragment that encodes a protein of interest is well known in the art. Preferably, the regenerated plants are self-pollinated to provide homozygous transgenic plants. Otherwise, pollen obtained from the regenerated plants is crossed to seed-grown plants of agronomically important lines. Conversely, pollen from plants of these important lines is used to pollinate regenerated plants. A transgenic plant of the present invention containing a desired polypeptide is cultivated using methods well known to one skilled in the art.

EXAMPLES

[0285] The present invention is further illustrated in the following Examples, in which parts and percentages are by weight and degrees are Celsius, unless otherwise stated. It should be understood that these Examples, while indicating preferred embodiments of the invention, are given by way of illustration only. From the above discussion and these Examples, one skilled in the art can ascertain the essential characteristics of this invention, and without departing from the spirit and scope thereof, can make various changes and modifications of the invention to adapt it to various usages and conditions. Thus, various modifications of the invention in addition to those shown and described herein will be apparent to those skilled in the art from the foregoing description. Such modifications are also intended to fall within the scope of the appended claims.

Example 1

Analysis of the rt1 Phenotype, Root Lodging and Root Penetration

[0286] The phenotype of rt1 homozygote plants is variable and depending on the growing conditions. As shown in FIG. 20 a-c, rt1 plants grown in the field have a very strong depletion of the root system and show a remarkably root lodging phenotype, while the same plants grown in greenhouse or hydroponic conditions show only a moderate alteration in root architecture. When rt1 plants are grown in the greenhouse using hard soil collected from the field, they still show a reduction in the root system (FIG. 20c), indicating that the rt1 locus is necessary for the penetration of the crown roots in the soil and the interaction of the roots with the soil. rt1 mutants are unable to push through hard soils.

[0287] Resistance to root lodging can be measured following mechanical perturbances under a variety of field conditions, such as vertical root pulling resistance (Beck et al. (1987); Crop Sci. 27:356-358), mechanized push (Kato et al. (1999) Maydica 44, 167-174) or a portable electronic design described by Fouere (Fouere et al. (1995), Agronomy J. 87: 1010-1024.

[0288] The design consists of a portable electronic device that measures horizontal pushing resistance on individual plants. The apparatus simultaneously records the angle of inclination and the resistance torque of the plant during an artificial pushing test. The device consists of a support, a force sensor, an angle sensor, and a control head. Data logging is possible by using a microprocessor-based system. Data may be transferred to a computer using an RS232 serial transfer protocol. The time required for the test in field conditions is approximately 1 min per plant. Preliminary results obtained on 14 maize genotypes grown in three field environments showed that genotypes susceptible to root lodging were characterized by low average values of their maximum resistance torque.

[0289] Root penetration into the soil can be measured, for example, by comparing seedling emergence and/or plant growth of RT1 transgenic and wild type plants on various soil types, e.g. tilled versus untilled soil. Untilled soils have higher resistance to root penetration and therefore reduce seedling emergence and subsequent plant growth. Soil types can be measured for example by using a cone penetrometer, measuring the tip resistance, sleeve friction and/or pore water pressure. The measurements can be made at different stages during plant growth and the difference in the rate of root penetration into the soil between a RT1 transgenic line and a nontransgenic (wild type) line can be recorded.

Example 2

Map-Based Cloning of RT1

[0290] The genetic data derived from the MaizeGDB (Maize Genetics and Genomics Database) showed that the rt1 mutant maps on chromosome 3 bin 04 of corn. Based on this information we retrieved from the public database several SSR primers and used them to genotype 88 rt1 plants derived from an F2 cross between the original rt1 line (unknown background, Jenkins M T, 1930) and the inbred line B73. Homozygous rt1/rt1 plants were scored as lodged plants when grown in the field for 30 days or more. DNA was extracted from those plants using standard molecular biology procedures. The public PCR-based DNA SSR marker UMC1908 (MaizeGDB) was found at 1.7 cM from the rt1 locus (3 recombinations on 88 individuals).

[0291] In order to fine map the rt1 mutation, two mapping populations and their corresponding corn seeds, segregating for the rt1 gene, were utilized. The first mapping populations consisted of 1500 BC2S1 plants derived by selfing a cross between a rt1 plant, derived from the above mentioned F2 population segregating for the rt1 locus, and the inbred line B73 (segregation ration 3:1). The second mapping populations consisted of 520 plants derived by backcrossing the above mentioned cross with the rt1 plant, parent of the cross (segregation ration 1:1).

[0292] To obtain plants that carry recombination near the rt1 locus, two sequence-based DNA markers, from the DuPont proprietary sequences of known map position were used.

[0293] Primers MZA8757-F81 (SEQ ID NO:1) and MZA8757-R593 (SEQ ID NO:2) were used to amplify and sequence the Mza8757 marker locus carrying a SNP A/C between the rt1 parent and the B73 parent.

[0294] Primers MZA15417-F132 (SEQ ID NO:3) and MZA15417-R607 (SEQ ID NO:4) were used to amplify and sequence the Mza 5417 marker locus carrying a SNP A/G between the rt1 parent and the B73 parent. Both markers reside on a physical BAC contig named 306 (Dupont Genomix database).

[0295] Only 62 plants showing a crossing over event between the two flanking markers (homozygotes at one marker and heterozygotes at the other marker) were kept and

selfed. The progeny of those plants was subsequently screened for the phenotype in order to confirm the position of the rt1 gene relatively to the flanking markers.

[0296] New CAPS (Cleaved Amplified Polymorphic Site) markers were designed using available physically placed MZA sequences and BAC-end sequences of the BACs constituting the region of contig 306 surrounded by markers Mza 15417 (left side) and Mza 8757 (right side).

[0297] CAPS primers were used in a PCR reaction containing 25 ng of DNA. CAPS marker amplifications were performed in a 25 ul PCR reaction using the Qiagen HotStart mix and 25 ng DNA. The thermal cycle conditions were: 95° C. 15 min (1 cycle), 94° C. 45 sec, 56° C. 45 sec, 72° C. 45 sec, (35 cycles) 72° C. 7 min.

[0298] 3 ul of the amplification product were used for a restriction digest (total volume of 15 ul) with the appropriate restriction enzyme. Restriction reaction was carried out at the appropriate temperature for one hour. Restricted amplification products were examined on 3% agarose gels.

[0299] CAPS marker b0541 (b0541 forward primer, SEQ ID NO:5 and b0541 reverse primer, SEQ ID NO:6) was designed based on the BAC-end sequence of clone BAC b0541.c13. This primer set amplifies a region of 250 bp, showing polymorphism between B73 and rt1 following restriction with the 6-cutter enzyme EcoRI.

[0300] CAPS marker b0461 (b0461 forward primer, SEQ ID NO:7 and b0461 reverse primer, SEQ ID NO:8): was designed based on the BAC-end sequence of clone BAC b0461.g10. This primer set amplifies a region of about 350 bp, showing polymorphism between B73 and rt1 following restriction with the 6-cutter enzyme NcoI.

[0301] By screening the 62 previously obtained recombinants with CAPS b0541, only 3 recombination breakpoints were found, while 2 recombinants were found on the other side, using the CAPS b0461.

[0302] BAC b0541.c13 and BAC b0461.g10 and are public BAC clones for which the available fingerprinting data show overlap.

Example 3

Identification of the RT1 Gene

[0303] In order to identify the RT1 gene that was mapped to the region comprising the two overlapping BAC clones, BAC b0541.c13 was sequenced. BAC DNA was nebulized using high-pressure nitrogen gas as described in Roe et al. 1996 (Roe et al. (1996) "DNA isolation and Sequencing" John Wiley and Sons, New York).

[0304] The estimated 165 Kb of sequence of BAC b0541. c13 was searched for the presence of open reading frames, and 4 regions, showing similarities to genes based on prediction models performed by FGENESH (Softberry, Inc. Mount Kisco, N.Y., USA) were identified.

[0305] In particular, one candidate gene showing homology to a "putative Ethylene Responsive Protein" was considered for further evaluation. The sequence of the gene derived from BAC b0541.c 3 (B73 genotype) is shown SEQ ID NO:9 (RT1 ethylene responsive gene 7800 bp) and in FIG. 1 Total RNA was extracted from developing maize roots from the B73 line and the rt1 line using a TRiazol® Reagent obtained from Life Technologies Inc., Rockville, Md., 20849 (GIBCO-BRL) that contains phenol and guanidine thiocyanate. RT-PCR was performed with cDNA that was synthesized with Superscript III (Invitrogen, Carlsbad, Calif.) reverse

transcriptase from 1 μg DNase treated total RNA. PCR was performed in a Perkin Elmer 9700 thermocycler using the GC-2 Advantage kit (BD Biosciences) and a PCR program of 94° C. for 3 min, followed by 27 cycles of 94° C. for 30 sec, 58° C. for 30 sec, 68° C. for 1 min, and a final step of 68° C. for 3 min. Primers designed based on the genomic sequence described in FIG. 1, SEQ ID NO:9 and specific to the 5' and 3' end of RT1 (RT1 3006F (SEQ ID NO:10) and RT17631R (SEQ ID NO:11), respectively, were used in the PCR reaction. Only the B73 wild type line generated a PCR product, indicating that the rt1 line is missing the RT1 mRNA, confirming that the lack of the RT1 transcript is responsible for the rt1 phenotype. The PCR product was cloned into the pPCR®II-Topo® nt vector (Invitrogen™) and sequenced to confirm identity. The B73 cDNA of RT1 is shown in SEQ ID NO:12. The RT1 amino acid sequence encoded by nucleotides 50 through 1382 (Stop codon) of SEQ ID NO:12 is shown in SEQ ID NO:13.

Example 4

Cloning the RT1 cDNA

[0306] Total RNA can be extracted from developing maize using a TRiazol® Reagent obtained from Life Technologies Inc., Rockville, Md., 20849 (GIBCO-BRL) that contains phenol and guanidine thiocyanate. Poly A mRNA can be purified from total RNA with mRNA Purification kits obtained from Amersham Pharmacia Biotech Inc., Piscataway, N.J., 08855, which consists of oligo (dT)-cellulose spin columns. To make the cDNA library, 5.5 ug of polyA RNA can be used for cDNA synthesis kits, which can be obtained from Stratagene, La Jolla, Calif., 92037. Superscript® reverse transcriptase can be obtained from Life Technologies Inc., Rockville, Md., 20849 (GIBCO-BRL). BRL cDNA Size Fraction Columns (GIBCO-BRL) can be used to fractionate the cDNA by size, fractions can be precipitated, resuspended and ligated with 1 ug of the Uni-ZAP XR vector. After ligation it can be packaged in Gigapack III Gold® packaging extract obtained from Stratagene, La Jolla, Calif., 92037. The unamplified library titer can be estimated. An appropriate amount can be used for amplification purposes to produce amplified cDNA.

[0307] Screening for the RT1 cDNA follows standard protocols well known to those skilled in the art (Ausubel et al. 1993, "Current Protocols in Molecular Biology" John Wiley & Sons, USA, or Sambrook et al. 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press). Briefly, 1.5×10^6 phage clones can be plated, then transferred to nylon membranes, which then will be subjected to hybridization with radioactively labeled RT1 probe. Positives are isolated and examined for their identity as RT1 cDNAs through PCR with RT1-specific primers. The longest cDNA clones that give positive results from the PCR reaction are isolated and sequenced.

Example 5

Ethylene Induction of the RT1 Gene

[0308] Promotor analysis of the RT1 gene via the PLACE database (Higo et al. (1999) *Nucleic Acids Res.* 27, 297-300) revealed two ERELEE4 ethylene response element (TTTGAATTT and TTTGAAAT) motives. In order to determine if the RT1 gene is induced by Ethylene treatment, B73 seedlings were germinated in paper rolls for 11 d in a phyto-

chamber at a 60% humidity, at 28° C., under a 16 h light, 8 h dark regime, then transferred to a 1.5×10^{-4} M Ethephon solution (Sigma Aldrich). Control plants were grown in distilled water. RNA was isolated from primary roots after 0, 1, 2.5 and 4 hours of Ethephon exposure and from control plants at the same time points. PCR reactions were performed with RT1 specific oligonucleotide primers (4405F, SEQ ID NO:14 and etr4Rnew, SEQ ID NO:15). Histone 2A (Genebank AAB04687) was used as control. FIG. **2** shows the induction of the RT1 gene at 2.5 hrs of treatment.

Example 6

Genetic Confirmation of the RT1 Gene

[0309] The genetic confirmation that the RT1 isolated nucleic acid fragment encodes the polypeptide responsible for altering root structure can be accomplished by transforming rt1 mutants with the isolated RT1 cloned sequence.

[0310] RT1 homologs from other crop species can also be tested in this system by obtaining full-gene sequences, ligation to an appropriate promoter, such as the RT1 promoter and complementing the maize rt1 mutant.

[0311] In order to confirm possible tissue-specific expression of the RT1 gene, the presence of the RT1 transcript in various tissues can be analyzed by RNA blot analysis and in situ hybridization.

[0312] One method for transforming DNA into cells of higher plants that is available to those skilled in the art is high-velocity ballistic bombardment using metal particles coated with the nucleic acid constructs of interest (see Klein et al. Nature (1987) (London) 327:70-73, and see U.S. Pat. No. 4,945,050). A Biolistic PDS-1000/He (BioRAD Laboratories, Hercules, Calif.) can be used for these complementation experiments. The particle bombardment technique can be used to transform the RT1 mutant with the cloned RT1 wild type sequence [nucleotides 50 through 1382 (Stop) of SEQ ID NO:12], encoding a functional RT1 protein.

[0313] The bacterial hygromycin B phosphotransferase (Hpt II) gene from *Streptomyces hygroscopicus* that confers resistance to the antibiotic hygromycin can be used as the selectable marker for the maize transformation. In the vector, pML18, the Hpt II gene can be engineered with the 35S promoter from Cauliflower Mosaic Virus and the termination and polyadenylation signals from the octopine synthase gene of *Agrobacterium tumefaciens*. pML18 was described in WO 97/47731, which was published on Dec. 18, 1997, the disclosure of which is hereby incorporated by reference.

[0314] Embryogenic maize callus cultures derived serve as source material for transformation experiments. This material can be generated by germinating sterile maize seeds on a callus initiation media (MS salts, Nitsch and Nitsch vitamins, 1.0 mg/l 2,4-D and 10 μM $AgNO_3$) in the dark at 27-28° C. Embryogenic callus proliferating from the scutellum of the embryos is then transferred to CM media (N6 salts, Nitsch and Nitsch vitamins, 1 mg/l 2,4-D, Chu et al., 1985, *Sci. Sinica* 18: 659-668). Callus cultures are maintained on CM by routine sub-culture at two week intervals and used for transformation within 10 weeks of initiation.

[0315] Callus can be prepared for transformation by subculturing 0.5-1.0 mm pieces approximately 1 mm apart, arranged in a circular area of about 4 cm in diameter, in the center of a circle of Whatman #541 paper placed on CM media. The plates with callus are incubated in the dark at 27-28° C. for 3-5 days. Prior to bombardment, the filters with

callus are transferred to CM supplemented with 0.25 M mannitol and 0.25 M sorbitol for 3 hr in the dark. The petri dish lids are then left ajar for 20-45 minutes in a sterile hood to allow moisture on tissue to dissipate.

[0316] Each genomic DNA fragment is co-precipitated with pML18 containing the selectable marker for maize transformation onto the surface of gold particles. To accomplish this, a total of 10 μg of DNA at a 2:1 ratio of trait: selectable marker DNAs are added to 50 μl aliquot of gold particles that are resuspended at a concentration of 60 mg ml^{-1}. Calcium chloride (50 μl of a 2.5 M solution) and spermidine (20 μl of a 0.1 M solution) are then added to the gold-DNA suspension as the tube was vortexed for 3 min. The gold particles are centrifuged in a microfuge for 1 sec and the supernatant removed. The gold particles are then washed twice with 1 ml of absolute ethanol and then resuspended in 50 ml of absolute ethanol and sonicated (bath sonicator) for one second to disperse the gold particles. The gold suspension is incubated at −70° C. for five minutes and sonicated (bath sonicator) if needed to disperse the particles. Six μl of the DNA-coated gold particles are then loaded onto mylar macrocarrier disks and the ethanol is allowed to evaporate.

[0317] At the end of the drying period, a petri dish containing the tissue is placed in the chamber of the PDS-1000/He. The air in the chamber is then evacuated to a vacuum of 28-29 inches Hg. The macrocarrier is accelerated with a helium shock wave using a rupture membrane that bursts when the He pressure in the shock tube reaches 1080-1100 psi. The tissue is placed approximately 8 cm from the stopping screen and the callus was bombarded two times. Two to four plates of tissue are bombarded in this way with the DNA-coated gold particles. Following bombardment, the callus tissue is transferred to CM media without supplemental sorbitol or mannitol.

[0318] Within 3-5 days after bombardment the callus tissue is transferred to SM media (CM medium containing 50 mg/l hygromycin). To accomplish this, callus tissue is transferred from plates to sterile 50 ml conical tubes and weighed. Molten top-agar at 40° C. is added using 2.5 ml of top agar/100 mg of callus. Callus clumps are broken into fragments of less than 2 mm diameter by repeated dispensing through a 10 ml pipette. Three ml aliquots of the callus suspension are plated onto fresh SM media and the plates are incubated in the dark for 4 weeks at 27-28° C. After 4 weeks, transgenic callus events are identified, transferred to fresh SM plates and grown for an additional 2 weeks in the dark at 27-28° C.

[0319] Growing callus can then be transferred to RM1 media (MS salts, Nitsch and Nitsch vitamins, 2% sucrose, 3% sorbitol, 0.4% gelrite +50 ppm hyg B) for 2 weeks in the dark at 25° C. After 2 weeks the callus can be transferred to RM2 media (MS salts, Nitsch and Nitsch vitamins, 3% sucrose, 0.4% gelrite+50 ppm hyg B) and placed under cool white light (~40 $\mu m^{-2} s^{-1}$) with a 12 hr photoperiod at 25° C. and 30-40% humidity. After 2-4 weeks in the light, callus can begin to organize, and form shoots. Shoots can be removed from surrounding callus/media and gently transferred to RM3 media (½×MS salts, Nitsch and Nitsch vitamins, 1% sucrose+50 ppm hygromycin B) in phytatrays (Sigma Chemical Co., St. Louis, Mo.) and incubation can be continued using the same conditions as described in the previous step.

[0320] Plants can then be transferred from RM3 to 4" pots containing Metro mix 350 after 2-3 weeks, when sufficient root and shoot growth has occurred. The seed obtained from

the transgenic plants can be examined for genetic complementation of the RT1 mutation with the wild-type genomic DNA containing the RT1 gene.

Example 7

Characterization of cDNA Clones Encoding RT1 Homologs

[0321] The BLASTX search using the EST sequences from clones listed in Table 1 revealed similarity of the polypeptides encoded by the ORF to proteins from rice and *Arabidopsis*. The nucleotide sequence encoding the closest polypeptide RT1 homolog from rice is shown in SEQ ID NO:16 and the corresponding amino acid sequence is set forth in NCBI General Identifier No: 115434026, SEQ ID NO:17). The nucleotide sequence encoding the closest polypeptide RT1 homolog from *Arabidopsis* is shown in SEQ ID NO:18 and the corresponding amino acid sequence is set forth in NCBI General Identifier No: 15217667, SEQ ID NO:19). The proteins from rice and *Arabidopsis* can be localized to Os01g01600 (TIGR) and At1g27660 (TAIR), respectively.

[0322] Shown in Table 1 and 2 are the literature and patent BLAST results, respectively, for individual ESTs ("EST"), the sequences of the entire cDNA inserts comprising the indicated cDNA clones ("FIS"), the sequences of contigs assembled from two or more ESTs ("Contig"), sequences of contigs assembled from an FIS and one or more ESTs ("Contig*"), or sequences encoding an entire protein derived from an FIS, a contig, or an FIS and P("CGS"). Also shown are the percent identities for sequences encoding RT1 and homologs thereof.

TABLE 1

BLAST Results (Literature) and Percent Identity
for Sequences Encoding RT1 and homologs thereof.

| | Sequence | |
| | B73-RT1 (SEQ ID NO: 12) cgs | B73-RT1 (SEQ ID NO: 13) protein |
Status		
BLAST pLOG Score to SEQ ID NO: 16)	49	N/A[1]
BLAST pLOG Score to NCBI GI No: 15217667 (SEQ ID NO: 18)	37	N/A
% identity to NCBI GI No: 115434026 (SEQ ID NO: 17)	N/A	39.7
% identity to NCBI GI No: 15217667 (SEQ ID NO: 19)	N/A	26.0

[1]N/A = non-applicable.

[0323] The BLASTX search using the Maize RT1 sequence (SEQ ID NO:12) revealed similarity to polypeptides homologous to RT1 from *Oryza sativa* (GI No. 115434026, SEQ ID NO:17) and to *Arabidopsis thaliana* (GI No. 15217667 SEQ ID NO:19) Shown in Table 1 and 2 are the BLAST results for individual ESTs ("EST"), the sequences of the entire cDNA inserts comprising the indicated cDNA clones ("FIS"), the sequences of contigs assembled from two or more EST, FIS or PCR sequences ("Contig"), or sequences encoding an entire or functional protein derived from an FIS or a contig ("CGS"):

TABLE 3

BLAST Results (patent) for Sequences Encoding
Polypeptides Homologous to RT1.

Sequence	Status	Reference	Blast pLog Score	% identity
B73-RT1 (SEQ ID NO: 13)	CGS	SEQ ID NO: 564 in WO2004031349-a2	38	29.3

[0324] Sequence alignments and percent identity calculations were performed using the Megalign program of the LASERGENE bioinformatics computing suite (DNASTAR Inc., Madison, Wis.). Multiple alignment of the sequences was performed using the Clustal method of alignment (Higgins and Sharp (1989) *CABIOS*. 5:151-153) with the default parameters (GAP PENALTY=10, GAP LENGTH PENALTY=10). Default parameters for pairwise alignments using the Clustal method were KTUPLE 1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5.

[0325] An additional maize RT1 homolog was found in the DuPont proprietary database. Maize clone cfp7n.pk6.13 (SEQ ID NO:20 encodes an RT1-like protein (SEQ ID NO:21) that has 47.2% identity to the B73 RT1 protein (SEQ ID NO:13) based on the Clustal method of alignment. An alignment of the maize RT1 protein and the rice, *Arabidospsis* and the maize homolog from clone cfp7n.pk6.i amino acid sequences (SEQ ID NO: 13; 17, 19, and 21) is shown jn FIGS. 3A-3B. Two sequence motifs (Motif I, SEQ ID NO:22 and Motif II, SEQ ID NO:23 in the alignment) are highly conserved in all four sequences and are shown underlined in the alignment.

Example 8

Knockout Analysis of the *Arabidopsis* RT1-Like Gene

[0326] In order to define the function of the RT1 gene in *Arabidopsis*, several Knockout lines, containing a T-DNA insertion in the At1g27660 locus (see Example 7) can be retrieved from the Salk Institute Genome Analysis Laboratory (SIGnAL) database. In particular, seeds from the two lines Salk_102156 and Salk_001968, segregating for the presence of the T-DNA insertion within the second intron of the gene, can be retrieved and planted. Seedlings can be genotyped using primers flanking the T-DNA insertions following the instructions provided by the database and roots of plants containing the T-DNA insertion in a homozygote state can be phenotyped using the software WinRHIZO® (Regent Instruments Inc). WinRHIZO® is an image analysis system specifically designed for root measurement which uses the contrast in pixels to distinguish the light root from the darker background (see also Example 21).

Example 9

Preparation of a Plant Expression Vector Containing the RT1 Gene or Homologs Thereof

[0327] Sequences homologous to the RT1 gene can be identified using sequence comparison algorithms such as BLAST (Basic Local Alignment Search Tool; Altschul et al., J. Mol. Biol. 215:403-410 (1993); see also the explanation of the BLAST algorithm on the world wide web site for the

22

National Center for Biotechnology Information at the National Library of Medicine of the National Institutes of Health). The RT1 gene (SEQ ID NO:12), or RT1-like genes, such as the one disclosed in SEQ ID NO:20, can be PCR-amplified by either of the following methods.

[0328] Method 1 (RNA-based): Based on the 5' and 3' sequence information for the protein-coding region of RT1 (extending from nts 50-1382 of SEQ ID NO:12) or a RT1 homolog (for example the sequence extending from nts 83-1540 of SEQ ID NO:20), gene-specific primers can be designed. RT-PCR can be used with plant RNA to obtain a nucleic acid fragment containing the RT1 protein-coding region flanked by attB1 (SEQ ID NO:24) and attB2 (SEQ ID NO:25) sequences. The primer may contain a consensus Kozak sequence (CAACA) upstream of the start codon.

[0329] Method 2 (DNA-based): Alternatively, the entire cDNA insert (containing 5' and 3' non-coding regions) of a clone encoding RT1 (SEQ ID NO:12 or a polypeptide homolog (such as the RT1 homolog encoded by SEQ ID NO:20), can be PCR amplified. Forward and reverse primers can be designed that contain either the attB1 sequence and vector-specific sequence that precedes the cDNA insert or the attB2 sequence and vector-specific sequence that follows the cDNA insert, respectively. For a cDNA insert cloned into the vector pBluescript SK+, the forward primer VC062 (SEQ ID NO:26) and the reverse primer VC063 (SEQ ID NO:27) can be used.

[0330] Methods 1 and 2 can be modified according to procedures known by one skilled in the art. For example, the primers of method 1 may contain restriction sites instead of attB1 and attB2 sites, for subsequent cloning of the PCR product into a vector containing attB1 and attB2 sites. Additionally, method 2 can involve amplification from a cDNA clone, a lambda clone, a BAC clone or genomic DNA.

[0331] A PCR product obtained by either method above can be combined with the Gateway® donor vector, such as PDONR™/Zeo (Invitrogen™, FIG. 5; SEQ ID NO:28) or PDONR™221 (Invitrogen™, FIG. 6; SEQ ID NO:29) using a BP Recombination Reaction. This process removes the bacteria lethal ccdB gene, as well as the chloramphenicol resistance gene (CAM) from the donor vectors and directionally clones the PCR product with flanking attB1 and attB2 sites to create an entry clone. Using the Invitrogen Gateway® Clonase™ technology, the RT1 or RT1-like gene from the entry clone can then be transferred to a suitable destination vector to obtain a plant expression vector for use with soy and corn, such as PHP27840 (FIG. 7; SEQ ID NO:30) or PHP23236 (FIG. 8; SEQ ID NO:31), respectively.

[0332] Alternatively a MultiSite Gateway® LR recombination reaction between multiple entry clones and a suitable destination vector can be performed to create an expression vector. An Example of this type of reaction is outlined in Example 14, which describes the construction of maize expression vectors for transformation of maize lines.

Example 10

Preparation of Soybean Expression Vectors and Transformation of Soybean with RT1 or Homologs Thereof

[0333] Soybean plants can be transformed to over-express the RT1 and homologs thereof, such as for example the RT1-like gene shown in SEQ ID NO:20 in order to examine the resulting phenotype.

[0334] The entry clones described in Example 9 can be used to directionally clone each gene into PHP27840 vector (FIG. 7, SEQ ID NO:30) such that expression of the gene is under control of the SCP1 promoter.

[0335] Soybean embryos may then be transformed with the expression vector comprising sequences encoding the instant polypeptides.

[0336] To induce somatic embryos, cotyledons, 3-5 mm in length dissected from surface sterilized, immature seeds of the soybean cultivar A2872, can be cultured in the light or dark at 26° C. on an appropriate agar medium for 6-10 weeks. Somatic embryos, which produce secondary embryos, are then excised and placed into a suitable liquid medium. After repeated selection for clusters of somatic embryos which multiply as early, globular staged embryos, the suspensions are maintained as described below.

[0337] Soybean embryogenic suspension cultures can be maintained in 35 mL liquid media on a rotary shaker, 150 rpm, at 26° C. with florescent lights on a 16:8 hour day/night schedule. Cultures are subcultured every two weeks by inoculating approximately 35 mg of tissue into 35 mL of liquid medium.

[0338] Soybean embryogenic suspension cultures may then be transformed by the method of particle gun bombardment (Klein et al. (1987) Nature (London) 327:70-73, U.S. Pat. No. 4,945,050). A DuPont Biolistic™ PDS1000/HE instrument (helium retrofit) can be used for these transformations.

[0339] A selectable marker gene which can be used to facilitate soybean transformation is a chimeric gene composed of the 35S promoter from cauliflower mosaic virus (Odell et al. (1985) Nature 313:810-812), the hygromycin phosphotransferase gene from plasmid pJR225 (from E. coli; Gritz et al. (1983) Gene 25:179-188) and the 3' region of the nopaline synthase gene from the T-DNA of the Ti plasmid of Agrobacterium tumefaciens. Another selectable marker gene which can be used to facilitate soybean transformation is an herbicide-resistant acetolactate synthase (ALS) gene from soybean or Arabidopsis. ALS is the first common enzyme in the biosynthesis of the branched-chain amino acids valine, leucine and isoleucine. Mutations in ALS have been identified that convey resistance to some or all of three classes of inhibitors of ALS (U.S. Pat. No. 5,013,659; the entire contents of which are herein incorporated by reference). Expression of the herbicide-resistant ALS gene can be under the control of a SAM synthetase promoter (U.S. Patent Application No. US-2003-0226166-A1; the entire contents of which are herein incorporated by reference).

[0340] To 50 μL of a 60 mg/mL 1 μm gold particle suspension is added (in order): 5 μL DNA (1 μg/μL), 20 μL spermidine (0.1 M), and 50 μL CaCl₂ (2.5 M). The particle preparation is then agitated for three minutes, spun in a microfuge for 10 seconds and the supernatant removed. The DNA-coated particles are then washed once in 400 μL 70% ethanol and resuspended in 40 μL of anhydrous ethanol. The DNA/particle suspension can be sonicated three times for one second each. Five μL of the DNA-coated gold particles are then loaded on each macro carrier disk.

[0341] Approximately 300-400 mg of a two-week-old suspension culture is placed in an empty 60×15 mm petri dish and the residual liquid removed from the tissue with a pipette. For each transformation experiment, approximately 5-10 plates of tissue are normally bombarded. Membrane rupture pressure is set at 1100 psi and the chamber is evacuated to a

vacuum of 28 inches mercury. The tissue is placed approximately 3.5 inches away from the retaining screen and bombarded three times. Following bombardment, the tissue can be divided in half and placed back into liquid and cultured as described above.

[0342] Five to seven days post bombardment, the liquid media may be exchanged with fresh media, and eleven to twelve days post bombardment with fresh media containing 50 mg/mL hygromycin. This selective media can be refreshed weekly. Seven to eight weeks post bombardment, green, transformed tissue may be observed growing from untransformed, necrotic embryogenic clusters. Isolated green tissue is removed and inoculated into individual flasks to generate new, clonally propagated, transformed embryogenic suspension cultures. Each new line may be treated as an independent transformation event. These suspensions can then be subcultured and maintained as clusters of immature embryos or regenerated into whole plants by maturation and germination of individual somatic embryos.

[0343] Enhanced root architecture can be measured in soybean by growing the plants in soil and wash the roots before analysis of the total root mass with the software WinRHIZO® (Regent Instruments Inc), an image analysis system specifically designed for root measurement. WinRHIZO® uses the contrast in pixels to distinguish the light root from the darker background.

[0344] Soybean plants transformed with the RT1 gene can then be assayed to study agronomic characteristics relative to control or reference plants. For example, nitrogen utilization efficacy, yield enhancement and/or stability under various environmental conditions (e.g. nitrogen limiting conditions, drought etc.).

Example 11

Transformation of Maize with the RT1 Gene and Homologs Thereof Using Particle Bombardment

[0345] Maize plants can be transformed to overexpress RT1 and RT1-like genes in order to examine the resulting phenotype.

[0346] The Gateway® entry clones described in Example 9 can be used to directionally clone each gene into a maize transformation vector. Expression of the gene in maize can be under control of a constitutive promoter such as the maize ubiquitin promoter (Christensen et al., *Plant Mol. Biol.* 12:619-632 (1989) and Christensen et al., *Plant Mol. Biol.* 18:675-689 (1992))

[0347] The recombinant DNA construct described above can then be introduced into maize cells by the following procedure. Immature maize embryos can be dissected from developing caryopses derived from crosses of the inbred maize lines H99 and LH132. The embryos are isolated ten to eleven days after pollination when they are 1.0 to 1.5 mm long. The embryos are then placed with the axis-side facing down and in contact with agarose-solidified N6 medium (Chu et al., *Sci. Sin. Peking* 18:659-668 (1975)). The embryos are kept in the dark at 27° C. Friable embryogenic callus consisting of undifferentiated masses of cells with somatic proembryoids and embryoids borne on suspensor structures proliferates from the scutellum of these immature embryos. The embryogenic callus isolated from the primary explant can be cultured on N6 medium and sub-cultured on this medium every two to three weeks.

[0348] The plasmid, p35S/Ac (obtained from Dr. Peter Eckes, Hoechst Ag, Frankfurt, Germany) may be used in transformation experiments in order to provide for a selectable marker. This plasmid contains the pat gene (see European Patent Publication 0 242 236) which encodes phosphinothricin acetyl transferase (PAT). The enzyme PAT confers resistance to herbicidal glutamine synthetase inhibitors such as phosphinothricin. The pat gene in p35S/Ac is under the control of the 35S promoter from cauliflower mosaic virus (Odell et al., *Nature* 313:810-812 (1985)) and the 3' region of the nopaline synthase gene from the T-DNA of the Ti plasmid of *Agrobacterium tumefaciens.*

[0349] The particle bombardment method (Klein et al., *Nature* 327:70-73 (1987)) may be used to transfer genes to the callus culture cells. According to this method, gold particles (1 μm in diameter) are coated with DNA using the following technique. Ten μg of plasmid DNAs are added to 50 μL of a suspension of gold particles (60 mg per mL). Calcium chloride (50 μL of a 2.5 M solution) and spermidine free base (20 μL of a 1.0 M solution) are added to the particles. The suspension is vortexed during the addition of these solutions. After ten minutes, the tubes are briefly centrifuged (5 sec at 15,000 rpm) and the supernatant removed. The particles are resuspended in 200 μL of absolute ethanol, centrifuged again and the supernatant removed. The ethanol rinse is performed again and the particles resuspended in a final volume of 30 μL of ethanol. An aliquot (5 μL) of the DNA-coated gold particles can be placed in the center of a Kapton™ flying disc (Bio-Rad Labs). The particles are then accelerated into the maize tissue with a Biolistic® PDS-1000/He (Bio-Rad Instruments, Hercules Calif.), using a helium pressure of 1000 psi, a gap distance of 0.5 cm and a flying distance of 1.0 cm.

[0350] For bombardment, the embryogenic tissue is placed on filter paper over agarose-solidified N6 medium. The tissue is arranged as a thin lawn and covered a circular area of about 5 cm in diameter. The petri dish containing the tissue can be placed in the chamber of the PDS-1000/He approximately 8 cm from the stopping screen. The air in the chamber is then evacuated to a vacuum of 28 inches of Hg. The macrocarrier is accelerated with a helium shock wave using a rupture membrane that bursts when the He pressure in the shock tube reaches 1000 psi.

[0351] Seven days after bombardment the tissue can be transferred to N6 medium that contains bialaphos (5 mg per liter) and lacks casein or proline. The tissue continues to grow slowly on this medium. After an additional two weeks the tissue can be transferred to fresh N6 medium containing bialophos. After six weeks, areas of about 1 cm in diameter of actively growing callus can be identified on some of the plates containing the bialaphos-supplemented medium. These calli may continue to grow when sub-cultured on the selective medium.

[0352] Plants can be regenerated from the transgenic callus by first transferring clusters of tissue to N6 medium supplemented with 0.2 mg per liter of 2,4-D. After two weeks the tissue can be transferred to regeneration medium (Fromm et al., *Bio/Technology* 8:833-839 (1990)).

[0353] Transgenic T0 plants can be regenerated and their phenotype determined following HTP procedures. T1 seed can be collected.

[0354] T1 plants can be grown and analyzed for phenotypic changes. The following parameters can be quantified using image analysis: plant area, volume, growth rate and color

analysis can be collected and quantified. Expression constructs that result in an alteration of root architecture compared to suitable control plants, can be considered evidence that the RT1 gene functions in maize to alter root architecture.

[0355] Furthermore, a recombinant DNA construct containing the RT1 gene can be introduced into an maize line either by direct transformation or introgression from a separately transformed line.

[0356] Transgenic plants, either inbred or hybrid, can undergo more vigorous field-based experiments to study yield enhancement, improved root penetration and/or resistance to root lodging under various environmental and soil conditions (e.g. variations in nutrient and water availability).

[0357] Subsequent yield analysis can also be done to determine whether plants that contain the RT1 gene have an improvement in yield performance, when compared to the control (or reference) plants that do not contain the RT1 gene. Plants containing the RT1 gene would have less yield loss relative to the control plants, preferably 50% less yield loss or would have increased yield relative to the control plants under varying environmental conditions.

Example 12

Electroporation of *Agrobacterium* LBA4404

[0358] Electroporation competent cells (40 μl), such as *Agrobacterium tumefaciens* LBA4404 (containing PHP10523, FIG. **9**, SEQ ID NO:32), are thawn on ice (20-30 min). PHP10523 contains VIR genes for T-DNA transfer, an *Agrobacterium* low copy number plasmid origin of replication, a tetracycline resistance gene, and a cos site for in vivo DNA biomolecular recombination. Meanwhile the electroporation cuvette is chilled on ice. The electroporator settings are adjusted to 2.1 kV. A DNA aliquot (0.5 μL JT (U.S. Pat. No. 7,087,812) parental DNA at a concentration of 0.2 μg-1.0 μg in low salt buffer or twice distilled H_2O) is mixed with the thawn *Agrobacterium* cells while still on ice. The mix is transferred to the bottom of electroporation cuvette and kept at rest on ice for 1-2 min. The cells are electroporated (Eppendorf electroporator 2510) by pushing "Pulse" button twice (ideally achieving a 4.0 msec pulse). Subsequently 0.5 ml 2×YT medium (or SOCmedium) are added to cuvette and transferred to a 15 ml Falcon tube. The cells are incubated at 28-30° C., 200-250 rpm for 3 h.

[0359] Aliquots of 250 μl are spread onto #30B (YM+50 μg/mL Spectinomycin) plates and incubated 3 days at 28-30° C. To increase the number of transformants one of two optional steps can be performed:

[0360] Option 1: overlay plates with 30 μl of 15 mg/ml Rifampicin. LBA4404 has a chromosomal resistance gene for Rifampicin. This additional selection eliminates some contaminating colonies observed when using poorer preparations of LBA4404 competent cells.

[0361] Option 2: Perform two replicates of the electroporation to compensate for poorer electrocompetent cells.

Identification of Transformants:

[0362] Four independent colonies are picked and streaked on AB minimal medium plus 50 mg/mL Spectinomycin plates (#12S medium) for isolation of single colonies. The plated are incubate at 28° C. for 2-3 days.

[0363] A single colony for each putative co-integrate is picked and inoculated with 4 ml #60A with 50 mg/l Spectinomycin. The mix is incubated for 24 h at 28° C. with shak-

ing. Plasmid DNA from 4 ml of culture is isolated using Qiagen Miniprep+optional PB wash. The DNA is eluted in 30 μl. Aliquots of 2 μl are used to electroporate 20 μl of DH10b+ 20 μl of ddH₂O as per above.

[0364] Optionally a 15 μl aliquot can be used to transform 75-100 μl of Invitrogen Library Efficiency DH5α. The cells are spread on LB medium plus 50 mg/mL Spectinomycin plates (#34T medium) and incubated at 37° C. overnight.

[0365] Three to four independent colonies are picked for each putative co-integrate and inoculated 4 ml of 2×YT (#60A) with 50 μg/ml Spectinomycin. The cells are incubated at 37° C. overnight with shaking.

[0366] Isolate plasmid DNA from 4 ml of culture using QIAprep® Miniprep with optional PB wash (elute in 50 μl). Use 8 μl for digestion with SalI (using JT parent and PHP10523 as controls).

[0367] Three more digestions using restriction enzymes BamHI, EcoRI, and HindIII are performed for 4 plasmids that represent 2 putative co-integrates with correct SalI digestion pattern (using parental DNA and PHP10523 as controls). Electronic gels are recommended for comparison.

[0368] Alternatively, for high throughput applications, such as described for Gaspe Bay Flint Derived Maize Lines (Examples 16-18), instead of evaluating the resulting co-integrate vectors by restriction analysis, three colonies can be simultaneously used for the infection step.

Example 13

Agrobacterium Mediated Transformation into Maize

[0369] Maize plants can be transformed to overexpress RT1 and RT1-like genes in order to examine the resulting phenotype.

[0370] *Agrobacterium*-mediated transformation of maize is performed essentially as described by Zhao et al., in *Meth. Mol. Biol.* 318:315-323 (2006) (see also Zhao et al., *Mol. Breed.* 8:323-333 (2001) and U.S. Pat. No. 5,981,840 issued Nov. 9, 1999, incorporated herein by reference). The transformation process involves bacterium innoculation, co-cultivation, resting, selection and plant regeneration.

1. Immature Embryo Preparation

[0371] Immature embryos are dissected from caryopses and placed in a 2 mL microtube containing 2 mL PHI-A medium.

2. *Agrobacterium* Infection and Co-Cultivation of Embryos

2.1 Infection Step

[0372] PHI-A medium is removed with 1 mL micropipettor and 1 mL *Agrobacterium* suspension is added. Tube is gently inverted to mix. The mixture is incubated for 5 min at room temperature.

2.2 Co-Culture Step

[0373] The *Agrobacterium* suspension is removed from the infection step with a 1 mL micropipettor. Using a sterile spatula the embryos are scraped from the tube and transferred to a plate of PHI-B medium in a 100×15 mm Petri dish. The embryos are oriented with the embryonic axis down on the surface of the medium. Plates with the embryos are cultured at 20° C., in darkness, for 3 days. L-Cysteine can be used in the co-cultivation phase. With the standard binary vector, the

co-cultivation medium supplied with 100-400 mg/L L-cysteine is critical for recovering stable transgenic events.

3. Selection of Putative Transgenic Events

[0374] To each plate of PHI-D medium in a 100×15 mm Petri dish, 10 embryos are transferred, maintaining orientation and the dishes are sealed with Parafilm. The plated are incubated in darkness at 28° C. Actively growing putative events, as pale yellow embryonic tissue are expected to be visible in 6-8 weeks. Embryos that produce no events may be brown and necrotic, and little friable tissue growth is evident. Putative transgenic embryonic tissue is subcultured to fresh PHI-D plates at 2-3 week intervals, depending on growth rate. The events are recorded.

4. Regeneration of T0 Plants

[0375] Embryonic tissue propagated on PHI-D medium is subcultured to PHI-E medium (somatic embryo maturation medium); in 100×25 mm Petri dishes and incubated at 28° C., in darkness, until somatic embryos mature, for about 10-18 days. Individual, matured somatic embryos with well-defined scutellum and coleoptile are transferred to PHI-F embryo germination medium and incubated at 28° C. in the light (about 80 μE from cool white or equivalent fluorescent lamps). In 7-10 days, regenerated plants, about 10 cm tall, are potted in horticultural mix and hardened-off using standard horticultural methods.

[0376] Media for Plant Transformation

    [0377]   1. PHI-A: 4 g/L CHU basal salts, 1.0 mL/L 1000× Eriksson's vitamin mix, 0.5 mg/L thiamin HCL, 1.5 mg/L 2,4-D, 0.69 g/L L-proline, 68.5 g/L sucrose, 36 g/L glucose, pH 5.2. Add 100 μM acetosyringone, filter-sterilized before using.

    [0378]   2. PHI-B: PHI-A without glucose, increased 2,4-D to 2 mg/L, reduced sucrose to 30 g/L and supplemented with 0.85 mg/L silver nitrate (filter-sterilized), 3.0 g/L gelrite, 100 μM acetosyringone (filter-sterilized), 5.8.

    [0379]   3. PHI-C: PHI-B without gelrite and acetosyringonee, reduced 2,4-D to 1.5 mg/L and supplemented with 8.0 g/L agar, 0.5 g/L Ms-morpholino ethane sulfonic acid (MES) buffer, 100 mg/L carbenicillin (filter-sterilized).

    [0380]   4. PHI-D: PHI-C supplemented with 3 mg/L bialaphos (filter-sterilized).

    [0381]   5. PHI-E: 4.3 g/L of Murashige and Skoog (MS) salts, (Gibco, BRL 11117-074), 0.5 mg/L nicotinic acid, 0.1 mg/L thiamine HCl, 0.5 mg/L pyridoxine HCl, 2.0 mg/L glycine, 0.1 g/L myo-inositol, 0.5 mg/L zeatin (Sigma, cat.no. Z-0164), 1 mg/L indole acetic acid (IAA), 26.4 μg/L abscisic acid (ABA), 60 g/L sucrose, 3 mg/L bialaphos (filter-sterilized), 100 mg/L carbenicillin (fileter-sterilized), 8 g/L agar, pH 5.6.

    [0382]   6. PHI-F: PHI-E without zeatin, IAA, ABA; sucrose reduced to 40 g/L; replacing agar with 1.5 g/L gelrite; pH 5.6.

[0383] Plants can be regenerated from the transgenic callus by first transferring clusters of tissue to N6 medium supplemented with 0.2 mg per liter of 2,4-D. After two weeks the tissue can be transferred to regeneration medium (Fromm et al. (1990) *Bio/Technology* 8:833-839).

[0384] Phenotypic analysis of transgenic T0 plants and T1 plants can be performed.

[0385] T1 plants can be analyzed for phenotypic changes. Using image analysis T1 plants can be analyzed for phenotypical changes in plant area, volume, growth rate and color analysis can be taken at multiple times during growth of the plants. Alteration in root architecture can be assayed as described In Example 21. Subsequent analysis of alterations in agronomic characteristics can be done to determine whether plants containing the RT1 or the RT1L gene have an improvement of at least one agronomic characteristic, when compared to the control (or reference) plants that do not contain RT1 or the RT1L gene. The alterations may also be studied under various environmental conditions.

Example 14

Construction of Maize Expression Vectors with the RT1 and RT1-Like Genes Using *Agrobacterium* Mediated Transformation

[0386] Maize expression vectors can be prepared with the RT1 (SEQ ID NO:12) and RT1-like genes (SEQ ID NO:20) under the control of the NAS2 (SEQ ID NO:33), GOS 2 (SEQ ID NO:34) or Ubiquitin (UBI1ZM; SEQ ID NO:35) promoter. PINII is the terminator (SEQ ID NO:36)

[0387] Using Invitrogen's™ Gateway® technology the entry clone, created as described in Example 9, containing the maize RT1 gene or maize RT-ike gene can be used in separate Gateway® LR reactions with:

[0388] 1) the constitutive maize GOS2 promoter entry clone PHP28408 (FIG. **10**, SEQ ID NO:37) and the PinII Terminator entry clone PHP20234 (FIG. **11**, SEQ ID NO:38), into the destination vector PHP28529 (FIG. **12**, SEQ ID NO:39).

[0389] 2) the root maize NAS2 promoter entry clone PHP22020 (FIG. **13**, SEQ ID NO:40) and the PinII Terminator entry clone PHP20234 (FIG. **11**, SEQ ID NO:38) into the destination vector PHP28529 (FIG. **12**, SEQ ID NO:39).

[0390] 3) the constitutive maize UBI1ZM promoter entry clone PHP23112 (FIG. **14**, SEQ ID NO:41) and the PinII Terminator entry clone PHP20234 (FIG. **11**, SEQ ID NO:38) into the destination vector PHP28529 (FIG. **12**, SEQ ID NO:39).

[0391] The destination vector PHP28529 adds to each of the final vectors also an:

    [0392]   1) RD29A promoter::yellow fluorescent protein:: PinII terminator cassette for *Arabidospis* seed sorting.

    [0393]   2) a Ubiquitin promoter::moPAT/red fluorescent protein fusion::PinII terminator cassette for transformation selection and *Z. mays* seed sorting.

[0394] In addition to the GOS2 or NAS2 promoter, other promoters such as, but not limited to the S2A and S2B promoter, the maize ROOTMET2 promoter, the maize Cyclo, the CR1BIO, the CRWAQ81 and the maize ZRP2.4447 are useful for directing expression of RT1 and RT1-like genes in maize. Furthermore, a variety of terminators, such as, but not limited to the PINII terminator, could be used to achieve expression of the gene of interest in maize.

Example 15

Transformation of Maize Lines with RT1 and RT1-Like Genes Using *Agrobacterium* Mediated Transformation

[0395] The final vectors (Example 14) can then electroporated separately into LBA4404 *Agrobacterium* containing

PHP10523 (FIG. 9; SEQ ID NO:32, Komari et al. Plant J 10:165-174 (1996), NCBI GI: 59797027) to create the co-integrate vectors for maize transformation. The co-integrate vectors are formed by recombination of the final vectors (maize expression vectors) with PHP10523, through the COS recombination sites contained on each vector. The co-integrate vectors contain in addition to the expression cassettes described in Example 14, also genes needed for the *Agrobacterium* strain and the *Agrobacterium* mediated transformation, (TET, TET, TRFA, ORI terminator, CTL, ORI V, VIR C1, VIR C2, VIR G, VIR B). Transformation into a maize line can be performed as described in Example 18.

Example 16

Preparation of the Destination Vectors PHP23236 and PHP29635 for Transformation of Gaspe Bay Flint Derived Maize Lines

[0396] Destination vector PHP23236 (FIG. 8, SEQ ID NO:31) was obtained by transformation of *Agrobacterium* strain LBA4404 containing plasmid PHP10523 (FIG. 9, SEQ ID NO:32) with plasmid PHP23235 (FIG. 15, SEQ ID NO:42) and isolation of the resulting co-integration product. Destination vector PHP23236, can be used in a recombination reaction with an entry clone as described in Example 9 to create a maize expression vector for transformation of Gaspe Bay Flint derived maize lines. Expression of the gene of interest is under control of the ubiquitin promoter (SEQ ID NO:35).

[0397] PHP29635 (FIG. 16, SEQ ID NO:43) was obtained by transformation of *Agrobacterium* strain LBA4404 containing plasmid PHP10523 with plasmid PHOXS2a-FRT87 (ni)m (FIG. 17, SEQ ID NO:44) and isolation of the resulting co-integration product. Destination vector PHP29635 can be used in a recombination reaction with an entry clone as described in Example 10 to create a maize expression vector for transformation of Gaspe Bay Flint derived maize lines. Expression of the gene of interest is under control of the S2A promoter (SEQ ID NO:45).

Example 17

Preparation of Plasmids Containing RT1 or RT1-Like Genes for Transformation of Gaspe Bay Flint Derived Maize Lines

[0398] Using Invitrogen's Gateway® Recombination technology, entry clones containing the RT1 or RT1-like genes can be created, as described in Example 10 and used to directionally clone each gene into destination vector PHP23236 (Example 16) for expression under the ubiquitin promoter or into destination vector PHP29635 (Example 16) for expression under the S2A promoter. Each of the expression vectors are T-DNA binary vectors for *Agrobacterium*-mediated transformation into corn.

[0399] Gaspe Bay Flint Derived Maize Lines can be transformed with the expression vectors as described in Example 18.

Example 18

Transformation of Gaspe Bay Flint Derived Maize Lines with RT1 and RT1-Like Genes

[0400] Maize plants can be transformed to over-express the RT1 and RT1-like genes, in order to examine the resulting phenotype.

[0401] Recipient Plants

[0402] Recipient plant cells can be from a uniform maize line having a short life cycle ("fast cycling"), a reduced size, and high transformation potential. Typical of these plant cells for maize are plant cells from any of the publicly available Gaspe Bay Flint (GBF) line varieties. One possible candidate plant line variety is the F1 hybrid of GBFxQTM (Quick Turnaround Maize, a publicly available form of Gaspe Bay Flint selected for growth under greenhouse conditions) disclosed in Tomes et al. U.S. Patent Application Publication No. 2003/0221212. Transgenic plants obtained from this line are of such a reduced size that they can be grown in four inch pots (¼ the space needed for a normal sized maize plant) and mature in less than 2.5 months. (Traditionally 3.5 months is required to obtain transgenic T0 seed once the transgenic plants are acclimated to the greenhouse.) Another suitable line is a double haploid line of GS3 (a highly transformable line) X Gaspe Flint. Yet another suitable line is a transformable elite inbred line carrying a transgene which causes early flowering, reduced stature, or both.

[0403] Transformation Protocol

[0404] Any suitable method may be used to introduce the transgenes into the maize cells, including and not limited to inoculation type procedures using *Agrobacterium* based vectors. Transformation may be performed on immature embryos of the recipient (target) plant.

[0405] Precision Growth and Plant Tracking

[0406] The event population of transgenic (T0) plants resulting from the transformed maize embryos is grown in a controlled greenhouse environment using a modified randomized block design to reduce or eliminate environmental error. A randomized block design is a plant layout in which the experimental plants are divided into groups (e.g., thirty plants per group), referred to as blocks, and each plant is randomly assigned a location with the block.

[0407] For a group of thirty plants, twenty-four transformed, experimental plants and six control plants (plants with a set phenotype) (collectively, a "replicate group") are placed in pots which are arranged in an array (a.k.a. a replicate group or block) on a table located inside a greenhouse. Each plant, control or experimental, is randomly assigned to a location with the block which is mapped to a unique, physical greenhouse location as well as to the replicate group. Multiple replicate groups of thirty plants each may be grown in the same greenhouse in a single experiment. The layout (arrangement) of the replicate groups should be determined to minimize space requirements as well as environmental effects within the greenhouse. Such a layout may be referred to as a compressed greenhouse layout.

[0408] An alternative to the addition of a specific control group is to identify those transgenic plants that do not express the gene of interest. A variety of techniques such as RT-PCR can be applied to quantitatively assess the expression level of the introduced gene. T0 plants that do not express the transgene can be compared to those which do.

[0409] Each plant in the event population is identified and tracked throughout the evaluation process, and the data gathered from that plant is automatically associated with that plant so that the gathered data can be associated with the transgene carried by the plant. For example, each plant container can have a machine readable label (such as a Universal Product Code (UPC) bar code) which includes information about the plant identity, which in turn is correlated to a green-

house location so that data obtained from the plant can be automatically associated with that plant.

[0410] Alternatively any efficient, machine readable, plant identification system can be used, such as two-dimensional matrix codes or even radio frequency identification tags (RFID) in which the data is received and interpreted by a radio frequency receiver/processor. See U.S. Published Patent Application No. 2004/0122592, incorporated herein by reference.

[0411] Phenotypic Analysis Using Three-Dimensional Imaging

[0412] Each greenhouse plant in the T0 event population, including any control plants, is analyzed for agronomic characteristics of interest, and the agronomic data for each plant is recorded or stored in a manner so that it is associated with the identifying data (see above) for that plant. Confirmation of a phenotype (gene effect) can be accomplished in the T1 generation with a similar experimental design to that described above.

[0413] The T0 plants are analyzed at the phenotypic level using quantitative, non-destructive imaging technology throughout the plant's entire greenhouse life cycle to assess the traits of interest. Preferably, a digital imaging analyzer is used for automatic multi-dimensional analyzing of total plants. The imaging may be done inside the greenhouse. Two camera systems, located at the top and side, and an apparatus to rotate the plant, are used to view and image plants from all sides. Images are acquired from the top, front and side of each plant. All three images together provide sufficient information to evaluate the biomass, size and morphology of each plant.

[0414] Due to the change in size of the plants from the time the first leaf appears from the soil to the time the plants are at the end of their development, the early stages of plant development are best documented with a higher magnification from the top. This may be accomplished by using a motorized zoom lens system that is fully controlled by the imaging software.

[0415] In a single imaging analysis operation, the following events occur: (1) the plant is conveyed inside the analyzer area, rotated 360 degrees so its machine readable label can be read, and left at rest until its leaves stop moving; (2) the side image is taken and entered into a database; (3) the plant is rotated 90 degrees and again left at rest until its leaves stop moving, and (4) the plant is transported out of the analyzer.

[0416] Plants are allowed at least six hours of darkness per twenty four hour period in order to have a normal day/night cycle.

[0417] Imaging Instrumentation

[0418] Any suitable imaging instrumentation may be used, including and not limited to light spectrum digital imaging instrumentation commercially available from LemnaTec GmbH of Wurselen, Germany. The images are taken and analyzed with a LemnaTec Scanalyzer HTS LT-0001-2 having a ½" IT Progressive Scan IEE CCD imaging device. The imaging cameras may be equipped with a motor zoom, motor aperture and motor focus. All camera settings may be made using LemnaTec software. Preferably, the instrumental variance of the imaging analyzer is less than about 5% for major components and less than about 10% for minor components.

[0419] Software

[0420] The imaging analysis system comprises a LemnaTec HTS Bonit software program for color and architecture analysis and a server database for storing data from about

500,000 analyses, including the analysis dates. The original images and the analyzed images are stored together to allow the user to do as much reanalyzing as desired. The database can be connected to the imaging hardware for automatic data collection and storage. A variety of commercially available software systems (e.g. Matlab, others) can be used for quantitative interpretation of the imaging data, and any of these software systems can be applied to the image data set.

[0421] Conveyor System

[0422] A conveyor system with a plant rotating device may be used to transport the plants to the imaging area and rotate them during imaging. For example, up to four plants, each with a maximum height of 1.5 m, are loaded onto cars that travel over the circulating conveyor system and through the imaging measurement area. In this case the total footprint of the unit (imaging analyzer and conveyor loop) is about 5 m×5 m.

[0423] The conveyor system can be enlarged to accommodate more plants at a time. The plants are transported along the conveyor loop to the imaging area and are analyzed for up to 50 seconds per plant. Three views of the plant are taken. The conveyor system, as well as the imaging equipment, should be capable of being used in greenhouse environmental conditions.

[0424] Illumination

[0425] Any suitable mode of illumination may be used for the image acquisition. For example, a top light above a black background can be used. Alternatively, a combination of top- and backlight using a white background can be used. The illuminated area should be housed to ensure constant illumination conditions. The housing should be longer than the measurement area so that constant light conditions prevail without requiring the opening and closing or doors. Alternatively, the illumination can be varied to cause excitation of either transgene (e.g., green fluorescent protein (GFP), red fluorescent protein (RFP)) or endogenous (e.g. Chlorophyll) fluorophores.

[0426] Biomass Estimation Based on Three-Dimensional Imaging

[0427] For best estimation of biomass the plant images should be taken from at least three axes, preferably the top and two side (sides 1 and 2) views. These images are then analyzed to separate the plant from the background, pot and pollen control bag (if applicable). The volume of the plant can be estimated by the calculation:

$$Volume(voxels) = \sqrt{TopArea(pixels)} \times \sqrt{Side1Area(pixels)} \times \sqrt{Side2Area(pixels)}$$

[0428] In the equation above the units of volume and area are "arbitrary units". Arbitrary units are entirely sufficient to detect gene effects on plant size and growth in this system because what is desired is to detect differences (both positive-larger and negative-smaller) from the experimental mean, or control mean. The arbitrary units of size (e.g. area) may be trivially converted to physical measurements by the addition of a physical reference to the imaging process. For instance, a physical reference of known area can be included in both top and side imaging processes. Based on the area of these physical references a conversion factor can be determined to allow conversion from pixels to a unit of area such as square centimeters (cm^2). The physical reference may or may not be an independent sample. For instance, the pot, with a known diameter and height, could serve as an adequate physical reference.

[0429] Color Classification

[0430] The imaging technology may also be used to determine plant color and to assign plant colors to various color classes. The assignment of image colors to color classes is an inherent feature of the LemnaTec software. With other image analysis software systems color classification may be determined by a variety of computational approaches.

[0431] For the determination of plant size and growth parameters, a useful classification scheme is to define a simple color scheme including two or three shades of green and, in addition, a color class for chlorosis, necrosis and bleaching, should these conditions occur. A background color class which includes non plant colors in the image (for example pot and soil colors) is also used and these pixels are specifically excluded from the determination of size. The plants are analyzed under controlled constant illumination so that any change within one plant over time, or between plants or different batches of plants (e.g. seasonal differences) can be quantified.

[0432] In addition to its usefulness in determining plant size growth, color classification can be used to assess other yield component traits. For these other yield component traits additional color classification schemes may be used. For instance, the trait known as "staygreen", which has been associated with improvements in yield, may be assessed by a color classification that separates shades of green from shades of yellow and brown (which are indicative of senescing tissues). By applying this color classification to images taken toward the end of the T0 or T1 plants' life cycle, plants that have increased amounts of green colors relative to yellow and brown colors (expressed, for instance, as Green/Yellow Ratio) may be identified. Plants with a significant difference in this Green/Yellow ratio can be identified as carrying transgenes which impact this important agronomic trait.

[0433] The skilled plant biologist will recognize that other plant colors arise which can indicate plant health or stress response (for instance anthocyanins), and that other color classification schemes can provide further measures of gene action in traits related to these responses.

[0434] Plant Architecture Analysis

[0435] Transgenes which modify plant architecture parameters may also be identified using the present invention, including such parameters as maximum height and width, internodal distances, angle between leaves and stem, number of leaves starting at nodes and leaf length. The LemnaTec system software may be used to determine plant architecture as follows. The plant is reduced to its main geometric architecture in a first imaging step and then, based on this image, parameterized identification of the different architecture parameters can be performed. Transgenes that modify any of these architecture parameters either singly or in combination can be identified by applying the statistical approaches previously described.

[0436] Pollen Shed Date

[0437] Pollen shed date is an important parameter to be analyzed in a transformed plant, and may be determined by the first appearance on the plant of an active male flower. To find the male flower object, the upper end of the stem is classified by color to detect yellow or violet anthers. This color classification analysis is then used to define an active flower, which in turn can be used to calculate pollen shed date.

[0438] Alternatively, pollen shed date and other easily visually detected plant attributes (e.g. pollination date, first silk date) can be recorded by the personnel responsible for performing plant care. To maximize data integrity and process efficiency this data is tracked by utilizing the same barcodes utilized by the LemnaTec light spectrum digital analyzing device. A computer with a barcode reader, a palm device, or a notebook PC may be used for ease of data capture recording time of observation, plant identifier, and the operator who captured the data.

[0439] Orientation of the Plants

[0440] Mature maize plants grown at densities approximating commercial planting often have a planar architecture. That is, the plant has a clearly discernable broad side, and a narrow side. The image of the plant from the broadside is determined. To each plant a well defined basic orientation is assigned to obtain the maximum difference between the broadside and edgewise images. The top image is used to determine the main axis of the plant, and an additional rotating device is used to turn the plant to the appropriate orientation prior to starting the main image acquisition.

Example 19

Screening of Gaspe Bay Flint Derived Maize Lines Under Nitrogen Limiting Conditions

[0441] Nitrogen utilization efficacy can be tested in the field by planting maize lines on nitrogen depleted soil or in the greenhouse using the experimental conditions as described herein. Transgenic plants will contain two or three doses of Gaspe Flint-3 with one dose of GS3 (GS3/(Gaspe-3)2× or GS3/(Gaspe-3)3×) and will segregate 1:1 for a dominant transgene. Plants will be planted in Turface, a commercial potting medium, and watered four times each day with 1 mM KNO_3 growth medium and with 2 mM KNO_3, or higher, growth medium (see FIG. 18). Control plants grown in 1 mM KNO_3 medium will be less green, produce less biomass and have a smaller ear at anthesis (see FIG. 19 for an illustration of sample data).

[0442] Statistics are used to decide if differences seen between treatments are really different. FIG. 19 illustrates one method which places letters after the values. Those values in the same column that have the same letter (not group of letters) following them are not significantly different. Using this method, if there are no letters following the values in a column, then there are no significant differences between any of the values in that column or, in other words, all the values in that column are equal.

[0443] Expression of a transgene will result in plants with improved plant growth in 1 mM KNO_3 when compared to a transgenic null. Thus biomass and greenness will be monitored during growth and compared to a transgenic null. Improvements in growth, greenness and ear size at anthesis will be indications of increased nitrogen tolerance.

Example 20

Yield Analysis of Maize Lines with RT1 or RT1-Like Genes

[0444] A recombinant DNA construct containing a RT1 or RT1-like Gene can be introduced into a maize line either by direct transformation or introgression from a separately transformed line.

[0445] Transgenic plants, either inbred or hybrid, can undergo more vigorous field-based experiments to study

yield enhancement and/or stability under various environmental conditions, such as variations in water and nutrient availability.

[0446] Subsequent yield analysis can be done to determine whether plants that contain the RT1 or RT1-like gene have an improvement in yield performance under various environmental conditions, when compared to the control plants that do not contain the RT1 or RT1-like gene. Reduction in yield can be measured for both. Plants containing the RT1 or RT1-like gene have less yield loss relative to the control plants, preferably 50% less yield loss.

Example 21

Assays to Determine Alterations of Root Architecture in Maize

[0447] Transgenic maize plants are assayed for changes in root architecture at seedling stage, flowering time or maturity. Assays to measure alterations of root architecture of maize plants include, but are not limited to the methods outlined below. To facilitate manual or automated assays of root architecture alterations, corn plants can be grown in clear pots.

[0448] 1) Root mass (dry weights). Plants are grown in Turface, a growth media that allows easy separation of roots. Oven-dried shoot and root tissues are weighed and a root/shoot ratio calculated.

[0449] 2) Levels of lateral root branching. The extent of lateral root branching (e.g. lateral root number, lateral root length) is determined by sub-sampling a complete root system, imaging with a flat-bed scanner or a digital camera and analyzing with WinRHIZO™ software (Regent Instruments Inc.).

[0450] 3) Root band width measurements. The root band is the band or mass of roots that forms at the bottom of greenhouse pots as the plants mature. The thickness of the root band is measured in mm at maturity as a rough estimate of root mass.

[0451] 4) Nodal root count. The number of crown roots coming off the upper nodes can be determined after separating the root from the support medium (e.g. potting mix). In addition the angle of crown roots and/or brace roots can be measured. Digital analysis of the nodal roots and amount of branching of nodal roots form another extension to the aforementioned manual method.

[0452] All data taken on root phenotype are subjected to statistical analysis, normally a t-test to compare the transgenic roots with that of non-transgenic sibling plants. One-way ANOVA may also be used in cases where multiple events and/or constructs are involved in the analysis.

Example 22

Screening of Gaspe Bay Flint Derived Maize Lines for Drought Tolerance

[0453] Transgenic Gaspe Bay Flint derived maize lines containing the RT1 or RT1-like gene can be screened for tolerance to drought stress in the following manner.

[0454] Transgenic maize plants are subjected to well-watered conditions (control) and to drought-stressed conditions. Transgenic maize plants are screened at the T1 stage or later.

[0455] Stress is imposed starting at 10 to 14 days after sowing (DAS) or 7 days after transplanting, and is continued through to silking. Pots are watered by an automated system

fitted to timers to provide watering at 25 or 50% of field capacity during the entire period of drought-stress treatment. The intensity and duration of this stress will allow identification of the impact on vegetative growth as well as on the anthesis-silking interval.

[0456] Potting mixture: A mixture of $\frac{1}{3}$ turface (Profile Products LLC, IL, USA), $\frac{1}{3}$ sand and $\frac{1}{3}$ SB300 (Sun Gro Horticulture, WA, USA) can be used. The SB300 can be replaced with Fafard Fine-Germ (Conrad Fafard, Inc., MA, USA) and the proportion of sand in the mixture can be reduced. Thus, a final potting mixture can be $\frac{3}{8}$ (37.5%) turface, $\frac{3}{8}$ (37.5%) Fafard and $\frac{1}{4}$ (25%) sand.

[0457] Field Capacity Determination: The weight of the soil mixture (w1) to be used in one S200 pot (minus the pot weight) is measured. If all components of the soil mix are not dry, the soil is dried at 100° C. to constant weight before determining w1. The soil in the pot is watered to full saturation and all the gravitational water is allowed to drain out. The weight of the soil (w2) after all gravitational water has seeped out (minus the pot weight) is determined. Field capacity is the weight of the water remaining in the soil obtained as w2-w1. It can be written as a percentage of the oven-dry soil weight.

[0458] Stress Treatment: During the early part of plant growth (10 DAS to 21 DAS), the well-watered control has a daily watering of 75% field capacity and the drought-stress treatment has a daily watering of 25% field capacity, both as a single daily dose at or around 10 AM. As the plants grow bigger, by 21 DAS, it will become necessary to increase the daily watering of the well-watered control to full field capacity and the drought stress treatment to 50% field capacity.

[0459] Nutrient Solution: A modified Hoagland's solution at $\frac{1}{16}$ dilution with tap water is used for irrigation.

TABLE 4

Preparation of 20 L of Modified Hoagland's
Solution Using the Following Recipe:

Component	Amount/20 L	
10X Micronutrient Solution	16	mL
KH$_2$PO$_4$ (MW: 136.02)	22	g
MgSO$_4$ (MW: 120.36)	77	g
KNO$_3$ (MW: 101.2)	129.5	g
Ca(NO$_3$)$_2$•4H$_2$0 (MW: 236.15)	151	g
NH$_4$NO$_3$ (MW: 80.04)	25.6	g
Sprint 330 (Iron chelate)	32	g

TABLE 5

Preparation of 1 L of 10X Micronutrient
Solution Using the Following Recipe:

Component	mg/L	Concentration
H$_3$BO$_3$	1854	30 mM
MnCl$_2$•4H$_2$0	1980	10 mM
ZnSO$_4$•7H$_2$0	2874	10 mM
CuSO$_4$•5H$_2$0	250	1 mM
H$_2$MoO$_4$•H$_2$0	242	1 mM

Fertilizer grade KNO$_3$ is used.

[0460] It is useful to add half a teaspoon of Osmocote (NPK 15:9:12) to the pot at the time of transplanting or after emergence (The Scotts Miracle-Gro Company, OH, USA).

[0461] Border plants: Place a row of border plants on bench-edges adjacent to the glass walls of the greenhouse or adjacent to other potential causes of microenvironment variability such as a cooler fan.

[0462] Automation: Watering can be done using PVC pipes with drilled holes to supply water to systematically positioned pots using a siphoning device. Irrigation scheduling can be done using timers.

[0463] Statistical analysis: Mean values for plant size, color and chlorophyll fluorescence recorded on transgenic events under different stress treatments will be exported to Spotfire (Spotfire, Inc., MA, USA). Treatment means will be evaluated for differences using Analysis of Variance.

[0464] Replications: Eight to ten individual plants are used per treatment per event.

[0465] Observations Made: Lemnatec measurements are made three times a week throughout growth to capture plant-growth rate. Leaf color determinations are made three times a week throughout the stress period using Lemnatec. Chlorophyll fluorescence is recorded as PhiPSII (which is indicative of the operating quantum efficiency of photosystem II photochemistry) and Fv'/Fm' (which is the maximum efficiency of photosystem II) two to four times during the experimental period, starting at 11 AM on the measurement days, using the Hansatech FMS2 instrument (LemnaTec GmbH, Wurselen, Germany). Measurements are started during the stress period at the beginning of visible drought stress symptoms, namely, leaf greying and the start of leaf rolling until the end of the experiment and measurements are recorded on the youngest most fully expanded leaf. The dates of tasseling and silking on individual plants are recorded, and the ASI is computed.

```
                         SEQUENCE LISTING


<160> NUMBER OF SEQ ID NOS: 45

<210> SEQ ID NO 1
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: artificial
<220> FEATURE:
<223> OTHER INFORMATION: primer

<400> SEQUENCE: 1

gcagacgcga caaagatgaa                                            20


<210> SEQ ID NO 2
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: artificial
<220> FEATURE:
<223> OTHER INFORMATION: primer

<400> SEQUENCE: 2

cgaatctcaa gctaatgagg a                                          21


<210> SEQ ID NO 3
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: artificial
<220> FEATURE:
<223> OTHER INFORMATION: primer

<400> SEQUENCE: 3

aggagcccag cgtgggcg                                              18


<210> SEQ ID NO 4
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: artificial
<220> FEATURE:
<223> OTHER INFORMATION: primer

<400> SEQUENCE: 4

caaaagaacc ccggagaagc ta                                         22


<210> SEQ ID NO 5
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: artificial
<220> FEATURE:
```

-continued

<223> OTHER INFORMATION: primer

<400> SEQUENCE: 5

ctatcccggt catccaatat gta   23

<210> SEQ ID NO 6
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: artificial
<220> FEATURE:
<223> OTHER INFORMATION: primer

<400> SEQUENCE: 6

gatctgccac tgggtattat gac   23

<210> SEQ ID NO 7
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: artificial
<220> FEATURE:
<223> OTHER INFORMATION: primer

<400> SEQUENCE: 7

gctgcagctt acattgttct ctc   23

<210> SEQ ID NO 8
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: artificial
<220> FEATURE:
<223> OTHER INFORMATION: primer

<400> SEQUENCE: 8

caacaccaga gcctctacag tg   22

<210> SEQ ID NO 9
<211> LENGTH: 7800
<212> TYPE: DNA
<213> ORGANISM: Zea mays

<400> SEQUENCE: 9

ggagatgttt cagtacatgc agagccttgg tgccgcatcc ggattccctg ctcaattcca   60

tactcttgtg agtatcaaaa ttttagtccc gcatgatata tattcatctg gtctcacaca   120

tgcaatatct tctctgtgca gagacaatcg gcggcatcga acaaccctca cgcttcaccc   180

agcccttcgt cgaaccagtc tagccgccca cctcgttgat gacattgtgt tggtgacatg   240

tttgagaaac ttgttgtgag atacttgttt gtgaaacttt tgtgagagac ctatgtttgt   300

gaaacttgtt tgatgtgttg atatttgtga gatatatggt gttggtgatg catgtgttga   360

ttctgtgata tatgtgatgc ctacgatatg tatcttttgt ttgtttggat ggaatagcaa   420

aaacaaataa aaagggtatt atggccactt tgccgagtgt aactctcggc aaagaaggca   480

cacctgggat ccggtaaagc ttctttgccg agtgttgtgg tccaggcact cagcaaaggg   540

actggcaaag gggcccatta gagccacctt tgccaagagc tcgtgcatca ggcactcgac   600

aaagaagtcc gctttgccga gtgcatatca tcgcactcgg cgaagggggct ggcaaaaggg   660

gtgcctagtg ccaggatagc agacacttga caaagactcc atcaccgtca cttggtatcg   720

cgagtttttct ttgccgagta ccaattgcac tcggtaaaga ctttatcaag tgtccgacaa   780

-continued

```
aaaagtactc ggcaaagaag ccgatctata gttcgcaaga cttctttgtc gagagtcaca      840

ctcggtaaag agtctgtcga gtgtttctg gactttgctg aagcacttgg aaaaacagtt        900

gtatccgtta gtgatgtagg gatggatttc ctaaagtctt gtaaaagaac gaataggcag       960

tgcacaactc aacaaacttg tatatacact acggagtact acactttag agctgcagta        1020

agaaaaagga gagagaaaca gtgaacgcaa tacatgggaa agctagggtc gggcggtacg       1080

gtaaggtatc tgccgtgccc atacctggca tctcatcaca atctcctggc caagcttgtc       1140

gatctagcct ctctctctct tttgtgcagt gagacccgac atgaatgggt gtggggctca       1200

tctgatcagt ccttgtgtta ctttgtttcc ttttcctttg aatttccggg gtttgcagaa       1260

ggacagccgg gtactgcatc gatgttgggc attctccagt tttgttcttg gtcttggagg       1320

aggaaaaggc ccagggacca acttttttcgc agccccccct tcttcctccc tttgcactgc      1380

tgctgccgtg aatcataca tcacaattac cgaggccaca gaaaagagtg ttttttttgct        1440

tgaggaaaag ggctagtccc tttattccaa ccacacacat gcatgcagaa aaagtcgatg       1500

gtggtggagt cgtacagtag tacaattaaa aactaactcg atcatatata tacatacttt       1560

gaaatgcaag tgccaggttc ttgagtactg tcctcagcta ctactacgta ccttgaagaa       1620

atttctagca agagacaatg gctaagaaag agaccatgca tggtgaaatt aaagagtgtt       1680

cttgacatga gaaacgtggt ttgctaactt taatcaagcc caaccaaagc tatcccaata       1740

tttaatggat gcatgggcta gccggccggg tgagagcaag gttaataata tagattataa       1800

tggcttttat gatgttgtca taccatgtat atagcaaact aaaagtccag tcatacaata       1860

actctaccat gatataagct atacatttaa tattaggact tatctttttt tttctcacac       1920

aaagtggctt agagtctatg agtaggccac ttttagctgg ctattagagc aagtttaata       1980

atataacgca tagtaagttt tataatgtta ccatgttaca tatagccaac aaaacatcta       2040

gtcatataat aactttcttg tgtatggtac aagttgtata tttaatattt attctactcc       2100

ttcctctcat aagttgtctt gtagtcagcc cacgtccagc tgactactag tgagcagccc       2160

gttttttctct cttttcttcgg tatggccata aatctgatgt gttatctctt gcagcttgtc     2220

tctgtcacct tattatactt gttcttagtg agcagtccgc ttctgtgtca ccgtaaacct       2280

gatagggcat gcacgtatta caatccgtcc acgtcatctt atcatgcttg ctctagtctt       2340

gttaaataga tctgtaataa accatatata gtccgttttt caagttggca tctttgctac       2400

ttccatttgg taccaacatg atctgacatc tctacgtcaa ctcgccagac agacaaatac       2460

ctattgcaga cgccaaagac aaacaataat tgacattact ttggccgttt tctgtagacc       2520

gttttctttt cccgagaaca atttgtacca gcagtgactg atcgatgagg cgttttagcc       2580

ctatatatgg cggctataat aatacacccg agccagaaaa gacttgtctc tttggtgcca       2640

tatgcatgtt tatggctacg ctagctttag ggcagagaga gagagagatc ctactagcta       2700

gcagaaatgt tagaatgata tgagagagag aacagagaga gcagagagag agagcttttc       2760

ttcttgcctt gaaaggagag gaacgggaac aaaggtaccc aatgctagca gatacataag       2820

gaggacgta cgggccgggt aaagcgacga agcctatgg atgggcatct ccatttctct         2880

aacccgatat gcctagctta gcttccattg gtggggcgt gtgtgttctt tgagagattc        2940

taaaagctcg tcaccgtaga gctagcagtg cgacctttat tttcttctcc ccctcgtggg      3000

ggaagtggca tcgatcgatt aagatcccct ttacaaccac catgatggca gcctcgtcgt       3060
```

-continued

```
cgtcgtcctt caacctcagt gatcatctcg ttcaagatga cctcccgtgg cctgcttctt    3120

cttcttcttc ctcccagctg ccatttgctc cttctcatcc tcctcacggc gccattattg    3180

gaggcagcca gcaatggtgg cagcaacctc aaatgctgta agcaccttgc agctctagct    3240

agcactagta tttccttatg tttacaatta ttgtcagggt caaaatttgt tctttgtgca    3300

ttgtgagtgt tcaatcatat tcccttcctc tctctctctc aaaggaaagt aattttggag    3360

tgcattcatg cattctttgt ttttatagta ctggattctt cctaggtaga atttggagtt    3420

ttcctctgca tatgcttgtt ctctctcccc cccccccccc ccagctgctc cctgttgtct    3480

ccatggagtc tagggaatcg tagctccggca aaatgacaac tttcggcaga aacagataaa    3540

cagcacacgg atgatgatgc aacttatgtg catgactcca tccaacaatt tgttggtttc    3600

ctccgcgact ccgaaaggtc gctgtgcatg cattttttctt ttttgtcatg gtcatcaatg    3660

gacggccggc atatatatac atacatacat agtgattcgt cctccacatc ttgtttaccc    3720

agcaggccag gccagccgga tgcatatgca tctatttgat tccttttact actgcagaaa    3780

ctgtgcagac cggcatccct ccgacgagct tgaggtcctc ctctcagcac aaggcagcca    3840

ccaccaccac caccatcctt cttcttcagt gatccatcct cagctaagct cgctgctaat    3900

gatgcaagac ctaggggttcc aatggagcag ctgcggcagc ttcgtggaca cctcctcgtc    3960

gtcgcccatg ccaccaccac tggtgaacgg ccagcagcac gacggccaca acaagatcaa    4020

ggagtgggac cgcccgctca ctgctaataa ctcatcaagg tcatcgtgtg cgagcacagc    4080

tatggcttgc cgagatgtcg tcctcgacgg cggcggcggc ggcggcctcc ctgccatggc    4140

cgccgtcgcc gccgatcttg acgacggcac cgtgctcccg cgagcgtca acgtcccccg    4200

gacgacgacg acgcgccaga ggccattccc ggcggcgccg ccacctctcc ccggcgacgc    4260

gttcgagata ctggcttcgt cgaggctatg caagaccctg ctgcttagcc aagcttcttc    4320

ttcggtactg ctccataatg gcacgacgat gcctttgttg agatcagagc atgtgcccta    4380

tggccctcct gctcctgctc atccccaagg accatccatc ggcaactaca agcaggtcag    4440

tcaaattaaa gtacagtgta gagcttgcgt cgttgcttga acagtacagt actctctctc    4500

tctctctctc tctcactctc agagaaggga aaggaaaagg atggatgcat gcatcaaacc    4560

gttagctgca tatgcagcag ctgctgtttt tttttttcttt ctttctttgg ctgaacagct    4620

gatgatactt atcctgcagt gtcttagaac tcaagcagta agtgtatagc cccatttgcc    4680

ctactgtctt gagacctcta gttaattgag ttcagacatt aaattgatag gagataactt    4740

aactaaaatc atttgaactc taaaagatag cataaaatct cctttttatga ttatatagtg    4800

tggcatgtaa actcaggtat atacatggta ttccatccat ccacctccag aaaactatat    4860

gcttgctgac ttgattagtc tcttgttagc caagccaaaa catgtttctt tcgtgagaga    4920

atgaactacg tgatggtata gttattgtgt atatacacgc atgtatatat atatatatat    4980

atatatatat atatatatat aggcgaagag agttcaagtc aataacagat taacagttac    5040

cagtagttct actctcccct atatatatgt gtgcatgttt tgctcatact ctctctatct    5100

atctatatgt ctatctgtcc gtgtccgtgt atatatgtgc atgcttcatt agagaaaata    5160

atggtcgatt tattttctaa gcccaaaaaa cagagtcaaa tatctatcta ctattataaa    5220

gcatcagatt cgatggtcgt cctacatcac gcgcgtaggt ctattcctct cccttttatca    5280

gtgtgggccg tgagaaacca cacaaggctc tattagtggg tcgttacctc gtagatgggc    5340
```

-continued

```
cgtgagtcaa gcctaacaac agaaagccgc atacagtgca acgcttccac atagtaccac    5400

tccactatct aaaggaggac gagctactgt gcgaactata aaaaagagg gtaaggctca     5460

ttttcaactc ataactttct aagacctttg gctaagataa aatgtaatat tttgttcttg    5520

tgtgtcgtgg caacacatgt gtattgtact agtctgtata tacacatagg atgtatgagt    5580

acctaagcta agagggtgtg tttgtttggg attataaact gttcagacta tataatccaa    5640

cagattttga actaataagt gttagttcaa aaatttattg gattatataa tctgggcaga    5700

ttataaaccc taaacaaaac atccctatgc cacctcacgt agctagcaac atgttacata    5760

tgtacagtat gtgatagaac aacatttgca tgccaaaagg cattgcgctt tcagtgtgag    5820

gaccattaca tttagctgtt attatgtgat cgagtagtac ggacaggggc gtacgattgt    5880

gttcaagctg ctaagaggtt acagctcatt gggtttgccg gtttgcgtgc cttgcctgct    5940

agagtaccag ctacacgcgc atgatatgat gtcctactcc tactcctaat ctcatgcatg    6000

catgtcctga tggatctgca gtgtgtgctg gcagccagtc tagtgcatgc tctctcgctc    6060

tccgatgaga tcctggccga agcaaaggta gagagagagc aggtcggagc agaagctagg    6120

cccagccggg ccgggcccat gctgtacaaa gaggaggagg aggagtacaa caggccggcc    6180

gggctctccc ttttgctctg catccattgc agcagccttt tagggttgga gagtgaccgc    6240

ccggccgcgc catctttttt ccgcccaata tcgtcgtacc gcccataccc tcctccggct    6300

ctggcctcgc tgcacttgca tcgtcttcac ttaactcctc gcaggctttg ttgtgtttgc    6360

agggaactaa taaagccgtt tgcagtgctg ctggcgccct gcccgtgccc tgtcaccacc    6420

atacgctaa ttcacacac cttattcata tatataatac accctccata tagatatata     6480

tagttgatga tatagaaaag agagcatgca tgggtgacag gctttgttct gttgaagaac    6540

gcaacgtgat cttaattctt ccctagctaa gatcagcggt cctccatcgt ttctgatgaa    6600

cacgcgtgcg tgcgtgcgcc tttgcagtgc atgcagatgg gcgcgtcttg tggaggccgc    6660

gcgcagcacg aagcggcatg tcaaagcatg gcggcggcag cagcacgccc agcactgcca    6720

gtggtgccgt cgtcgaagaa gcctcgagtc gagcagcggt cctccaccat cgtcccttct    6780

ttcaaggtac tcgtatactc taggctgggt ctcgtgccgt ctcgatcgtt tcgtcatttc    6840

gatcaacgtt aattgacagt aacaccaacc taatatccaa tcatatatat atagtactaa    6900

ttctgttccc tgattaggta aggaaagaga aattaggggga ccggatcgct gctcttcagc    6960

agctcgtctc accgtttggc aaggtacccg gccgtcgtca cgattattca gcattattag    7020

ggttttggag ggcaggcagg attacatata ttattaactc tctctctata tatataaata    7080

tatacctgac actaacacta ctatgtatgt gcagactgat accgcatctg tgctcatgga    7140

ggccatcggt tacatcaaat ttcttcaaga tcaagtcgag gtaaccaact agctagctgc    7200

ttatcagcca cacttggtcc atgagttcca gacatagcca cgctcgccaa gaaatactgt    7260

ttctttcttt ccatggcaag catctgctgg taggattttt ttcctcctga cagcagttct    7320

catgatcacc ctcaccctca cctgcagact ctgagccgcc cctacctgaa atcctccaga    7380

agtaagaagc ccaggccgac accgcgggta acaccatacg tctaaatcag gctcggtttc    7440

aaaatctggg ggaaactttc atgttaaatc tgatgtttgt gcacgtcgcg ttgcgtgtga    7500

atcagggttc ttcaaacgcg agcgctgggg agctggagga aacgaggcct gacctccgca    7560

gcagagggct ctgcctggtg cctctctcgt gcacgtccta cgtgaccaac gagaacggag    7620
```

-continued

tctgggttcc accgaatttt agggcaaact aaagtcatgt cagcgctact agagacacag   7680

cacaacagtg taccaagaac tactaatcca cggttcatct cagttcatcc acggttaata   7740

ccagttctcc gggccaaaag acacccacca tctacctctt gtgttgtctc tctttgccac   7800

<210> SEQ ID NO 10
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: artificial
<220> FEATURE:
<223> OTHER INFORMATION: primer

<400> SEQUENCE: 10

tggcatcgat cgattaagat cccct   25

<210> SEQ ID NO 11
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: artificial
<220> FEATURE:
<223> OTHER INFORMATION: primer

<400> SEQUENCE: 11

tggaacccag actccgttct cgt   23

<210> SEQ ID NO 12
<211> LENGTH: 1385
<212> TYPE: DNA
<213> ORGANISM: Zea mays

<400> SEQUENCE: 12

ctcgtggggg aagtggcatc gatcgattaa gatccccttt acaaccacca tgatggcagc   60

ctcgtcgtcg tcgtccttca acctcagtga tcatctcgtt caagatgacc tcccgtggcc   120

tgcttcttct tcttcttcct cccagctgcc atttgctcct tctcatcctc ctcacggcgc   180

cattattgga ggcagccagc aatggtggca gcaacctcaa atgctaaact gtgcagaccg   240

gcatccctcc gacgagcttg aggtcctcct ctcagcacaa ggcagccacc accaccacca   300

ccatccttct tcttcagtga tccatcctca gctaagctcg ctgctaatga tgcaagacct   360

agggttccaa tggagcagct gcggcagctt cgtggacacc tcctcgtcgt cgcccatgcc   420

accaccactg gtgaacggcc agcagcacga cggccacaac aagatcaagg agtgggaccg   480

cccgctcact gctaataact catcaaggtc atcgtgtgcg agcacagcta tggcttgccg   540

agatgtcgtc ctcgacggcg gcggcggcgg cggcctccct gccatggccg ccgtcgccgc   600

cgatcttgac gacggcaccg tgctcccggc gagcgtcaac gtcccccgga cgacgacgac   660

gcgccagagg ccattcccgg cggcgccgcc acctctcccc ggcgacgcgt tcgagatact   720

ggcttcgtcg aggctatgca agaccctgct gcttagccaa gcttcttctt cggtactgct   780

ccataatggc acgacgatgc ctttgttgag atcagagcat gtgccctatg gccctcctgc   840

tcctgctcat ccccaaggac catccatcgg caactacaag cagtgcatgc agatgggcgc   900

gtcttgtgga ggccgcgcgc agcacgaagc ggcatgtcaa agcatggcgg cggcagcagc   960

acgcccagca ctgccagtgg tgccgtcgtc gaagaagcct cgagtcgagc agcggtcctc   1020

caccatcgtc ccttctttca aggtaaggaa agagaaatta ggggaccgga tcgctgctct   1080

tcagcagctc gtctcaccgt ttggcaagac tgataccgca tctgtgctca tggaggccat   1140

-continued

```
cggttacatc aaatttcttc aagatcaagt cgagactctg agccgcccct acctgaaatc    1200

ctccagaagt aagaagccca ggccgacacc gcggggttct tcaaacgcga gcgctgggga    1260

gctggaggaa acgaggcctg acctccgcag cagagggctc tgcctggtgc ctctctcgtg    1320

cacgtcctac gtgaccaacg agaacggagt ctgggttcca ccgaatttta gggcaaacta    1380

aagtc                                                                1385
```

```
<210> SEQ ID NO 13
<211> LENGTH: 443
<212> TYPE: PRT
<213> ORGANISM: Zea mays

<400> SEQUENCE: 13

Met Met Ala Ala Ser Ser Ser Ser Ser Phe Asn Leu Ser Asp His Leu
1               5                   10                  15

Val Gln Asp Asp Leu Pro Trp Pro Ala Ser Ser Ser Ser Ser Ser Gln
            20                  25                  30

Leu Pro Phe Ala Pro Ser His Pro Pro His Gly Ala Ile Ile Gly Gly
        35                  40                  45

Ser Gln Gln Trp Trp Gln Gln Pro Gln Met Leu Asn Cys Ala Asp Arg
    50                  55                  60

His Pro Ser Asp Glu Leu Glu Val Leu Leu Ser Ala Gln Gly Ser His
65                  70                  75                  80

His His His His Pro Ser Ser Ser Val Ile His Pro Gln Leu Ser
                85                  90                  95

Ser Leu Leu Met Met Gln Asp Leu Gly Phe Gln Trp Ser Ser Cys Gly
            100                 105                 110

Ser Phe Val Asp Thr Ser Ser Ser Pro Met Pro Pro Leu Val
        115                 120                 125

Asn Gly Gln Gln His Asp Gly His Asn Lys Ile Lys Glu Trp Asp Arg
    130                 135                 140

Pro Leu Thr Ala Asn Asn Ser Ser Arg Ser Ser Cys Ala Ser Thr Ala
145                 150                 155                 160

Met Ala Cys Arg Asp Val Val Leu Asp Gly Gly Gly Gly Gly Leu
                165                 170                 175

Pro Ala Met Ala Ala Val Ala Ala Asp Leu Asp Asp Gly Thr Val Leu
            180                 185                 190

Pro Ala Ser Val Asn Val Pro Arg Thr Thr Thr Thr Arg Gln Arg Pro
        195                 200                 205

Phe Pro Ala Ala Pro Pro Leu Pro Gly Asp Ala Phe Glu Ile Leu
    210                 215                 220

Ala Ser Ser Arg Leu Cys Lys Thr Leu Leu Leu Ser Gln Ala Ser Ser
225                 230                 235                 240

Ser Val Leu Leu His Asn Gly Thr Thr Met Pro Leu Leu Arg Ser Glu
                245                 250                 255

His Val Pro Tyr Gly Pro Pro Ala Pro Ala His Pro Gln Gly Pro Ser
            260                 265                 270

Ile Gly Asn Tyr Lys Gln Cys Met Gln Met Gly Ala Ser Cys Gly Gly
        275                 280                 285

Arg Ala Gln His Glu Ala Ala Cys Gln Ser Met Ala Ala Ala Ala
    290                 295                 300

Arg Pro Ala Leu Pro Val Val Pro Ser Ser Lys Lys Pro Arg Val Glu
```

-continued

```
305                310                315                320

Gln Arg Ser Ser Thr Ile Val Pro Ser Phe Lys Val Arg Lys Glu Lys
                325                330                335

Leu Gly Asp Arg Ile Ala Ala Leu Gln Gln Leu Val Ser Pro Phe Gly
            340                345                350

Lys Thr Asp Thr Ala Ser Val Leu Met Glu Ala Ile Gly Tyr Ile Lys
        355                360                365

Phe Leu Gln Asp Gln Val Glu Thr Leu Ser Arg Pro Tyr Leu Lys Ser
    370                375                380

Ser Arg Ser Lys Lys Pro Arg Pro Thr Pro Arg Gly Ser Ser Asn Ala
385                390                395                400

Ser Ala Gly Glu Leu Glu Glu Thr Arg Pro Asp Leu Arg Ser Arg Gly
                405                410                415

Leu Cys Leu Val Pro Leu Ser Cys Thr Ser Tyr Val Thr Asn Glu Asn
            420                425                430

Gly Val Trp Val Pro Pro Asn Phe Arg Ala Asn
        435                440


<210> SEQ ID NO 14
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: artificial
<220> FEATURE:
<223> OTHER INFORMATION: primer

<400> SEQUENCE: 14

ccaaggacca tccatcggca act                                      23


<210> SEQ ID NO 15
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: artificial
<220> FEATURE:
<223> OTHER INFORMATION: primer

<400> SEQUENCE: 15

actccgttct cgttggtcac                                          20


<210> SEQ ID NO 16
<211> LENGTH: 1167
<212> TYPE: DNA
<213> ORGANISM: Oryza sativa

<400> SEQUENCE: 16

atgatggcct cctcctcttc ctcctccagc ctctgtgata tcatctctt gcaagatgac   60

ctgatcccat ggccttccat gcccttttgct cctgctccca acacctttgg attgaaccac  120

caatggagcc agcctccaat gctaagcagc agtactgatc agttaagctc atatgagctt  180

gaatcactcc aatcagttga aagtcagctg gcagctgctg ctcctccgac attaagccc    240

catctgcagg ctcatcaact cagcacggtg ctaatgatgc aggagctagg gttccaatgg   300

agcagctgcg ctgcgccagc agatcaacat agtattgcaa gctccacgaa taataatagt   360

aatgtgatga tgaacgagga ggagctacgg ccacgtcctg atcaatccct aatcagcaat   420

ccaaggtcat gcagcgccac cacgcttctt cctccgccgc atctacatct ggacggcgct  480

gtgctcccga gcatcaacgt ctcgcggctg cagaagctgg ccgccggcga tgagccgctg  540
```

-continued

```
cagatatgtt gcaagaggca ggctgctgct gctgcagtgg tcggccatag cagcataaga    600

gatgagcatg tgccctgccc ctatgctggc cctcctgctc atctaattca aggaccatcc    660

aacaccctgc agatgaagag gaatacgaat gcagcagctc aagggagagg tggccgccat    720

gggagttcaa cggagcacag atcatcaaca gcgctgccgc catcgtcaaa gaagcccgg     780

ctcgagtcac actcctcctc aatgcttcct tcattcaagg taaggaaaga gaaactaggg    840

gatagaattg ccgcccttca gcaactggtc tcaccctttg gtaagactga taccgcatcg    900

gtactcatgg aggccatcgg ttacatcaaa tttcttcagg accaagttga gactttgagc    960

ggtccatacc tgagatcatc caagaacagc aagaaattag cctgcagggc agcacaacaa    1020

caacggaagg ggacgtcaaa tggaggggac gcagcagcaa agcttgacct ccgaagtaga    1080

ggactctgcc tggtgcccct gtcgtgcacc tcctacgtca ccaacgagaa tggggtttgg    1140

cctccaccca atttcagagg aaactga                                       1167
```

```
<210> SEQ ID NO 17
<211> LENGTH: 388
<212> TYPE: PRT
<213> ORGANISM: Oryza sativa

<400> SEQUENCE: 17

Met Met Ala Ser Ser Ser Ser Ser Ser Ser Leu Cys Asp Asn His Leu
1               5                   10                  15

Leu Gln Asp Asp Leu Ile Pro Trp Pro Ser Met Pro Phe Ala Pro Ala
                20                  25                  30

Pro Asn Thr Phe Gly Leu Asn His Gln Trp Ser Gln Pro Pro Met Leu
            35                  40                  45

Ser Ser Ser Thr Asp Gln Leu Ser Ser Tyr Glu Leu Glu Ser Leu Gln
        50                  55                  60

Ser Val Glu Ser Gln Leu Ala Ala Ala Ala Pro Pro Thr Leu Ser Pro
65                  70                  75                  80

His Leu Gln Ala His Gln Leu Ser Thr Val Leu Met Met Gln Glu Leu
                85                  90                  95

Gly Phe Gln Trp Ser Ser Cys Ala Ala Pro Ala Asp Gln His Ser Ile
            100                 105                 110

Ala Ser Ser Thr Asn Asn Asn Ser Asn Val Met Met Asn Glu Glu Glu
        115                 120                 125

Leu Arg Pro Arg Pro Asp Gln Ser Leu Ile Ser Asn Pro Arg Ser Cys
    130                 135                 140

Ser Ala Thr Thr Leu Leu Pro Pro Pro His Leu His Leu Asp Gly Ala
145                 150                 155                 160

Val Leu Pro Ser Ile Asn Val Ser Arg Leu Gln Lys Leu Ala Ala Gly
                165                 170                 175

Asp Glu Pro Leu Gln Ile Cys Cys Lys Arg Gln Ala Ala Ala Ala
            180                 185                 190

Val Val Gly His Ser Ser Ile Arg Asp Glu His Val Pro Cys Pro Tyr
            195                 200                 205

Ala Gly Pro Pro Ala His Leu Ile Gln Gly Pro Ser Asn Thr Leu Gln
    210                 215                 220

Met Lys Arg Asn Thr Asn Ala Ala Ala Gln Gly Arg Gly Gly Arg His
225                 230                 235                 240

Gly Ser Ser Thr Glu His Arg Ser Ser Thr Ala Leu Pro Pro Ser Ser
```

-continued

```
              245               250               255
Lys Lys Pro Arg Leu Glu Ser His Ser Ser Ser Met Leu Pro Ser Phe
        260               265               270

Lys Val Arg Lys Glu Lys Leu Gly Asp Arg Ile Ala Ala Leu Gln Gln
        275               280               285

Leu Val Ser Pro Phe Gly Lys Thr Asp Thr Ala Ser Val Leu Met Glu
    290               295               300

Ala Ile Gly Tyr Ile Lys Phe Leu Gln Asp Gln Val Glu Thr Leu Ser
305               310               315               320

Gly Pro Tyr Leu Arg Ser Ser Lys Asn Ser Lys Lys Leu Ala Cys Arg
            325               330               335

Ala Ala Gln Gln Gln Arg Lys Gly Thr Ser Asn Gly Gly Asp Ala Ala
        340               345               350

Ala Lys Leu Asp Leu Arg Ser Arg Gly Leu Cys Leu Val Pro Leu Ser
        355               360               365

Cys Thr Ser Tyr Val Thr Asn Glu Asn Gly Val Trp Pro Pro Pro Asn
    370               375               380

Phe Arg Gly Asn
385
```

<210> SEQ ID NO 18
<211> LENGTH: 1362
<212> TYPE: DNA
<213> ORGANISM: Arabidopsis thaliana

<400> SEQUENCE: 18

```
atggactctg ctaatctcca tcagcttcaa gatcaattac agcttgtggg gtcttcttca      60

tcttcttctt ccttagacaa taactctgac ccttcttgct atggagcttc atctgcccat     120

caatggagcc aggaggtat ttctttgaat agtgtgagct tgagtcataa ttataacaat     180

gagatgttaa acacaagagc tcacaacaac aacaacaaca caacacaag tgaatgtatg     240

agtctctcca gcatccacaa tcactccttg atccaacaac aagactttcc tttacaatgg     300

cctcatgacc aatcttcata tcaacatcat gaaggacttc tcaagatcaa agaagagctt     360

tcctcatcaa ctatctcaga ccatcaagaa ggcatatcca agttcacaga catgttaaat     420

agtccagtga taacaaacta tttgaagatc aatgaacata aggactacac tgagaagctt     480

cttctcaaga gtatgtcttc tggattcccg atcaatggag actatggtag cagccttccc     540

tcttcttctt cttcctcttc accttcgtct cagtcgcata gaggcaactt cagtcagatt     600

tacccaagcg taaacatatc gagtttgagc gaatctcgga agatgagcat ggacgacatg     660

agtaacatct caagaccatt tgatataaac atgcaggttt ttgatggaag attgtttgaa     720

ggaaatgtat tagttcctcc ttttaacgct caagagatta gtagtcttgg atgagcaga     780

ggaagccttc cttctttggg cctccctttt catcatcatc tgcagcaaac acttccccac     840

ctttcttctt ccctactca tcaaatggaa atgttcagca atgaacctca aacaagtgaa     900

gggaagagagc ataacttctt gatggcaaca aaagcaggag aaaatgcttc caagaaaccg     960

cgcgtggaat cacgctcctc ttgcccaccc ttcaaggtga ggaaagaaaa gttaggagac    1020

agaatagcag ctctgcagca gttggtttca ccctttggga gacagatac agcatctgtg    1080

ttaatggaag caattggata catcaaattc ctacagagcc agtcgagac tttaagcgtc    1140

ccctacatga gagcatctag gaaccgaccc ggaaaagcct cccagctggt ctcacaatca    1200
```

-continued

caagaagggg atgaggaaga gacgagagat cttagaagcc gtgggctatg tctagtgccg   1260

ttatcatgca tgacttatgt taccggagat ggtggggatg gaggaggcgg tgttggtact   1320

ggttttttggc caacgccacc tggttttggt ggcggaactt ag   1362

<210> SEQ ID NO 19
<211> LENGTH: 453
<212> TYPE: PRT
<213> ORGANISM: Arabidopsis thaliana

<400> SEQUENCE: 19

Met Asp Ser Ala Asn Leu His Gln Leu Gln Asp Gln Leu Gln Leu Val
1   5   10   15

Gly Ser Ser Ser Ser Ser Ser Ser Leu Asp Asn Asn Ser Asp Pro Ser
            20   25   30

Cys Tyr Gly Ala Ser Ser Ala His Gln Trp Ser Pro Gly Gly Ile Ser
        35   40   45

Leu Asn Ser Val Ser Leu Ser His Asn Tyr Asn Asn Glu Met Leu Asn
    50   55   60

Thr Arg Ala His Asn Asn Asn Asn Asn Asn Thr Ser Glu Cys Met
65   70   75   80

Ser Leu Ser Ser Ile His Asn His Ser Leu Ile Gln Gln Gln Asp Phe
                85   90   95

Pro Leu Gln Trp Pro His Asp Gln Ser Ser Tyr Gln His His Glu Gly
            100   105   110

Leu Leu Lys Ile Lys Glu Glu Leu Ser Ser Ser Thr Ile Ser Asp His
            115   120   125

Gln Glu Gly Ile Ser Lys Phe Thr Asp Met Leu Asn Ser Pro Val Ile
    130   135   140

Thr Asn Tyr Leu Lys Ile Asn Glu His Lys Asp Tyr Thr Glu Lys Leu
145   150   155   160

Leu Leu Lys Ser Met Ser Ser Gly Phe Pro Ile Asn Gly Asp Tyr Gly
                165   170   175

Ser Ser Leu Pro Ser Ser Ser Ser Ser Ser Pro Ser Ser Gln Ser
            180   185   190

His Arg Gly Asn Phe Ser Gln Ile Tyr Pro Ser Val Asn Ile Ser Ser
        195   200   205

Leu Ser Glu Ser Arg Lys Met Ser Met Asp Asp Met Ser Asn Ile Ser
    210   215   220

Arg Pro Phe Asp Ile Asn Met Gln Val Phe Asp Gly Arg Leu Phe Glu
225   230   235   240

Gly Asn Val Leu Val Pro Pro Phe Asn Ala Gln Glu Ile Ser Ser Leu
                245   250   255

Gly Met Ser Arg Gly Ser Leu Pro Ser Phe Gly Leu Pro Phe His His
            260   265   270

His Leu Gln Gln Thr Leu Pro His Leu Ser Ser Ser Pro Thr His Gln
        275   280   285

Met Glu Met Phe Ser Asn Glu Pro Gln Thr Ser Glu Gly Lys Arg His
    290   295   300

Asn Phe Leu Met Ala Thr Lys Ala Gly Glu Asn Ala Ser Lys Lys Pro
305   310   315   320

Arg Val Glu Ser Arg Ser Ser Cys Pro Pro Phe Lys Val Arg Lys Glu

-continued

```
              325             330             335
Lys Leu Gly Asp Arg Ile Ala Ala Leu Gln Gln Leu Val Ser Pro Phe
            340             345             350
Gly Lys Thr Asp Thr Ala Ser Val Leu Met Glu Ala Ile Gly Tyr Ile
        355             360             365
Lys Phe Leu Gln Ser Gln Ile Glu Thr Leu Ser Val Pro Tyr Met Arg
    370             375             380
Ala Ser Arg Asn Arg Pro Gly Lys Ala Ser Gln Leu Val Ser Gln Ser
385             390             395             400
Gln Glu Gly Asp Glu Glu Glu Thr Arg Asp Leu Arg Ser Arg Gly Leu
            405             410             415
Cys Leu Val Pro Leu Ser Cys Met Thr Tyr Val Thr Gly Asp Gly Gly
            420             425             430
Asp Gly Gly Gly Gly Val Gly Thr Gly Phe Trp Pro Thr Pro Pro Gly
        435             440             445
Phe Gly Gly Gly Thr
    450
```

```
<210> SEQ ID NO 20
<211> LENGTH: 1759
<212> TYPE: DNA
<213> ORGANISM: Zea Mays

<400> SEQUENCE: 20

cgctactagc tagtgcgact ttttctttct ctccccctct tggaagtagc atcgatcgat       60

taaggccgcc tttacaacca ccatgatggc ggcttcgtca tcgtcctcca acctcagtga      120

tcacctcgcg caagatgacc tcctcccatg gccttcttct tcttcctcct ccctgccatt      180

tgctcctgct gctccccatc acagcgccat tattggaggc aacagccagc aacctctagc      240

gctaaactgc gctgaccggc atcactccga cgagcttgag gtgctcctct cagcacaagg      300

cagccaccac caccagagcc acagccacag gcacagccac cctgcttctc cggtgatccc      360

tccgcagctc agcagctctc tgctaacgat gcaagaccta gggttccaat ggagcaactg      420

cggcggcggc ttcctagaca cctcaccgtc cacgccgccg acgaacgggc agctgcagca      480

cgatggccac gacgacgaca agatcgaggc ggagggcacg cttattacta actcgtcaag      540

gccatcgtgt gcgggcacag ccacgatagc cgccagctgc cacgacgtcg tcctccacgg      600

cggcggcggc ggcggcaccg tgctcccgac cataaacatc tcgctgacga cgacgacgac      660

gacgcacacg caggcgcagc agcagcagcg accgttcctg gcggcgccgc cgcctctccc      720

gggcgacgcg ttcgagatac tactggcctc gtcgaggctg tgcaagaccc tgctgcttag      780

ccaagctgct gcttcttctt cttcttcttc ggtgctgctc cacgacgacg gcacgatcag      840

cacggtgcct tcgctgagat cggagtcgga gcatgtggcc tatggctatg cggcggccc      900

tcctcctcct gctcctgccg ctgctcatcg tccccgagga cgacgaccat ccgttgacaa      960

ctacaagcag cagatgggcg ccgccgcctg tgccgcggga gcaggccgcc agtggagcgc     1020

ggagcacgga gggcgagccg ccgcctgtga agaaagggcg gcagcggcac cagcacgccg     1080

cccggtggtg gtgccatgga agaagcctcg actcgagcag cgctccggct ccggctccgg     1140

ctccggctcc tccaccacca tcctcccttc tttcgaggta cggaaagaga agttaggggа     1200

ccggatcgct gccctccagc agctcgtctc gccattcggc aagaccgata ccgcgtctgt     1260
```

-continued

```
gctcatggag gccattggct acatcaagtt cctccaagac caagtcgaga ctttgagccg    1320

cccctacttg aagtcgtcca gaagtagcaa gaagcccagg ccgacgcacc agcggggggtg    1380

ttggaacgcc agcgctgcag gggaggagca ggagcaggag gagacgacga cgaggaggcg    1440

gcggcctgac ctccgcagca gagggctctg cctcgtgccc ctctcctgca cgtcctacgt    1500

caccaacgag aacgggcct gggttccttc ccaatttttaa ttttagggga gactactagc    1560

tatagagatc agaggagcag agacaacact gtaccacgaa ctaatccacg gttaatatca    1620

cctcacatct ctgctctgct catgttgttg tctgtgtctc tttgccttgg cccttggccg    1680

taaataccag tcagtttcta aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa    1740

aaaaaaaaaa aaaaaaaaa                                                 1759
```

```
<210> SEQ ID NO 21
<211> LENGTH: 485
<212> TYPE: PRT
<213> ORGANISM: Zea mays

<400> SEQUENCE: 21

Met Met Ala Ala Ser Ser Ser Ser Ser Asn Leu Ser Asp His Leu Ala
1               5                   10                  15

Gln Asp Asp Leu Leu Pro Trp Pro Ser Ser Ser Ser Ser Ser Leu Pro
            20                  25                  30

Phe Ala Pro Ala Ala Pro His His Ser Ala Ile Ile Gly Gly Asn Ser
        35                  40                  45

Gln Gln Pro Leu Ala Leu Asn Cys Ala Asp Arg His His Ser Asp Glu
    50                  55                  60

Leu Glu Val Leu Leu Ser Ala Gln Gly Ser His His His Gln Ser His
65                  70                  75                  80

Ser His Arg His Ser His Pro Ala Ser Pro Val Ile Pro Pro Gln Leu
                85                  90                  95

Ser Ser Ser Leu Leu Thr Met Gln Asp Leu Gly Phe Gln Trp Ser Asn
            100                 105                 110

Cys Gly Gly Gly Phe Leu Asp Thr Ser Pro Ser Thr Pro Pro Thr Asn
        115                 120                 125

Gly Gln Leu Gln His Asp Gly His Asp Asp Asp Lys Ile Glu Ala Glu
    130                 135                 140

Gly Thr Leu Ile Thr Asn Ser Ser Arg Pro Ser Cys Ala Gly Thr Ala
145                 150                 155                 160

Thr Ile Ala Ala Ser Cys His Asp Val Val Leu His Gly Gly Gly Gly
                165                 170                 175

Gly Gly Thr Val Leu Pro Thr Ile Asn Ile Ser Leu Thr Thr Thr Thr
            180                 185                 190

Thr Thr His Thr Gln Ala Gln Gln Gln Arg Pro Phe Leu Ala Ala
            195                 200                 205

Pro Pro Pro Leu Pro Gly Asp Ala Phe Glu Ile Leu Leu Ala Ser Ser
    210                 215                 220

Arg Leu Cys Lys Thr Leu Leu Leu Ser Gln Ala Ala Ala Ser Ser Ser
225                 230                 235                 240

Ser Ser Ser Val Leu Leu His Asp Asp Gly Thr Ile Ser Thr Val Pro
                245                 250                 255

Ser Leu Arg Ser Glu Ser Glu His Val Ala Tyr Gly Tyr Gly Gly Gly
            260                 265                 270
```

```
Pro Pro Pro Pro Ala Pro Ala Ala Ala His Arg Pro Arg Gly Arg Arg
        275             280             285

Pro Ser Val Asp Asn Tyr Lys Gln Gln Met Gly Ala Ala Ala Cys Ala
        290             295             300

Ala Gly Ala Gly Arg Gln Trp Ser Ala Glu His Gly Gly Arg Ala Ala
305             310             315             320

Ala Cys Glu Glu Arg Ala Ala Ala Ala Pro Ala Arg Arg Pro Val Val
            325             330             335

Val Pro Trp Lys Lys Pro Arg Leu Glu Gln Arg Ser Gly Ser Gly Ser
        340             345             350

Gly Ser Gly Ser Ser Thr Thr Ile Leu Pro Ser Phe Glu Val Arg Lys
        355             360             365

Glu Lys Leu Gly Asp Arg Ile Ala Ala Leu Gln Gln Leu Val Ser Pro
        370             375             380

Phe Gly Lys Thr Asp Thr Ala Ser Val Leu Met Glu Ala Ile Gly Tyr
385             390             395             400

Ile Lys Phe Leu Gln Asp Gln Val Glu Thr Leu Ser Arg Pro Tyr Leu
            405             410             415

Lys Ser Ser Arg Ser Ser Lys Lys Pro Arg Pro Thr His Gln Arg Gly
            420             425             430

Cys Trp Asn Ala Ser Ala Ala Gly Glu Glu Gln Glu Gln Glu Glu Thr
        435             440             445

Thr Thr Arg Arg Arg Arg Pro Asp Leu Arg Ser Arg Gly Leu Cys Leu
        450             455             460

Val Pro Leu Ser Cys Thr Ser Tyr Val Thr Asn Glu Asn Gly Ala Trp
465             470             475             480

Val Pro Ser Gln Phe
                485
```

```
<210> SEQ ID NO 22
<211> LENGTH: 47
<212> TYPE: PRT
<213> ORGANISM: artificial
<220> FEATURE:
<223> OTHER INFORMATION: Motif I

<400> SEQUENCE: 22

Val Arg Lys Glu Lys Leu Gly Asp Arg Ile Ala Ala Leu Gln Gln Leu
1               5               10              15

Val Ser Pro Phe Gly Lys Thr Asp Thr Ala Ser Val Leu Met Glu Ala
                20              25              30

Ile Gly Tyr Ile Lys Phe Leu Gln Ser Gln Ile Glu Thr Leu Ser
            35              40              45
```

```
    <210> SEQ ID NO 23
<211> LENGTH: 14
<212> TYPE: PRT
<213> ORGANISM: artificial
<220> FEATURE:
<223> OTHER INFORMATION: Motif II

<400> SEQUENCE: 23

Asp Leu Arg Ser Arg Gly Leu Cys Leu Val Pro Leu Ser Cys
1               5               10
```

```
<210> SEQ ID NO 24
<211> LENGTH: 29
<212> TYPE: DNA
<213> ORGANISM: artificial
<220> FEATURE:
<223> OTHER INFORMATION: attB1 recombination site

<400> SEQUENCE: 24

ggggacaagt ttgtacaaaa aagcaggct                                          29


<210> SEQ ID NO 25
<211> LENGTH: 29
<212> TYPE: DNA
<213> ORGANISM: artificial
<220> FEATURE:
<223> OTHER INFORMATION: attB2 recombination site

<400> SEQUENCE: 25

ggggaccact ttgtacaaga aagctgggt                                          29


<210> SEQ ID NO 26
<211> LENGTH: 54
<212> TYPE: DNA
<213> ORGANISM: artificial
<220> FEATURE:
<223> OTHER INFORMATION: primer

<400> SEQUENCE: 26

ttaaacaagt ttgtacaaaa aagcaggctg caattaaccc tcactaaagg gaac            54


<210> SEQ ID NO 27
<211> LENGTH: 53
<212> TYPE: DNA
<213> ORGANISM: artificial
<220> FEATURE:
<223> OTHER INFORMATION: primer

<400> SEQUENCE: 27

ttaaaccact ttgtacaaga aagctgggtg cgtaatacga ctcactatag ggc             53


<210> SEQ ID NO 28
<211> LENGTH: 4291
<212> TYPE: DNA
<213> ORGANISM: artificial
<220> FEATURE:
<223> OTHER INFORMATION: vector

<400> SEQUENCE: 28

ctttcctgcg ttatcccctg attctgtgga taaccgtatt accgcctttg agtgagctga      60

taccgctcgc cgcagccgaa cgaccgagcg cagcgagtca gtgagcgagg aagcggaaga     120

gcgcccaata cgcaaaccgc ctctccccgc gcgttggccg attcattaat gcagctggca     180

cgacaggttt cccgactgga aagcgggcag tgagcgcaac gcaattaata cgcgtaccgc     240

tagccaggaa gagtttgtag aaacgcaaaa aggccatccg tcaggatggc cttctgctta     300

gtttgatgcc tggcagttta tggcgggcgt cctgcccgcc accctccggg ccgttgcttc     360

acaacgttca atccgctccc ggcggatttg tcctactca  ggagagcgtt caccgacaaa     420

caacagataa aacgaaaggc ccagtcttcc gactgagcct ttcgtttat  ttgatgcctg     480

gcagttccct actctcgcgt taacgctagc atggatgttt tcccagtcac gacgttgtaa     540

aacgacggcc agtcttaagc tcgggcccca ataatgatt  ttattttgac tgatagtgac     600
```

-continued

```
ctgttcgttg caacacattg atgagcaatg cttttttata atgccaactt tgtacaaaaa      660

agctgaacga gaaacgtaaa atgatataaa tatcaatata ttaaattaga ttttgcataa      720

aaaacagact acataatact gtaaaacaca acatatccag tcactatgaa tcaactactt      780

agatggtatt agtgacctgt agtcgaccga cagccttcca aatgttcttc gggtgatgct      840

gccaacttag tcgaccgaca gccttccaaa tgttcttctc aaacggaatc gtcgtatcca      900

gcctactcgc tattgtcctc aatgccgtat taaatcataa aaagaaataa gaaaagagg      960

tgcgagcctc tttttttgtgt gacaaaataa aaacatctac ctattcatat acgctagtgt     1020

catagtcctg aaaatcatct gcatcaagaa caatttcaca actcttatac ttttctctta     1080

caagtcgttc ggcttcatct ggattttcag cctctatact tactaaacgt gataaagttt     1140

ctgtaatttc tactgtatcg acctgcagac tggctgtgta aagggagcc tgacatttat     1200

attccccaga acatcaggtt aatggcgttt ttgatgtcat tttcgcggtg gctgagatca     1260

gccacttctt ccccgataac ggagaccggc acactggcca tatcggtggt catcatgcgc     1320

cagctttcat ccccgatatg caccaccggg taaagttcac gggagacttt atctgacagc     1380

agacgtgcac tggccagggg gatcaccatc cgtcgcccgg gcgtgtcaat aatatcactc     1440

tgtacatcca caaacagacg ataacggctc tctcttttat aggtgtaaac cttaaactgc     1500

atttcaccag ccctgttct cgtcagcaaa agagccgttc atttcaataa accgggcgac     1560

ctcagccatc ccttcctgat tttccgcttt ccagcgttcg gcacgcagac gacgggcttc     1620

attctgcatg gttgtgctta ccagaccgga gatattgaca tcatatatgc cttgagcaac     1680

tgatagctgt cgctgtcaac tgtcactgta atacgctgct tcatagcata cctcttttg     1740

acatacttcg ggtatacata tcagtatata ttcttatacc gcaaaaatca gcgcgcaaat     1800

acgcatactg ttatctggct tttagtaagc cggatccacg cggcgtttac gcccgccct     1860

gccactcatc gcagtactgt tgtaattcat taagcattct gccgacatgg aagccatcac     1920

agacggcatg atgaacctga atcgccagcg gcatcagcac cttgtcgcct tgcgtataat     1980

atttgcccat ggtgaaaacg ggggcgaaga agttgtccat attggccacg tttaaatcaa     2040

aactggtgaa actcacccag ggattggctg agacgaaaaa catattctca ataaaccctt     2100

tagggaaata ggccaggttt tcaccgtaac acgccacatc ttgcgaatat atgtgtagaa     2160

actgccggaa atcgtcgtgg tattcactcc agagcgatga aaacgtttca gtttgctcat     2220

ggaaaacggt gtaacaaggg tgaacactat cccatatcac cagctcaccg tctttcattg     2280

ccatacggaa ttccggatga gcattcatca ggcgggcaag aatgtgaata aaggccggat     2340

aaaacttgtg cttattttc tttacggtct ttaaaaaggc cgtaatatcc agctgaacgg     2400

tctggttata ggtacattga caactgact gaaatgcctc aaaatgttct ttacgatgcc     2460

attgggatat atcaacggtg gtatatccag tgatttttttt ctccatttta gcttccttag     2520

ctcctgaaaa tctcgataac tcaaaaaata cgcccggtag tgatcttatt tcattatggt     2580

gaaagttgga acctcttacg tgccgatcaa cgtctcattt tcgccaaaag ttggcccagg     2640

gcttcccggt atcaacaggg acaccaggat ttatttattc tgcgaagtga tcttccgtca     2700

caggtatta ttcggcgcaa agtgcgtcgg gtgatgctgc caacttagtc gactacaggt     2760

cactaatacc atctaagtag ttgattcata tgactggat atgttgtgtt ttacagtatt     2820

atgtagtctg tttttttatgc aaaatctaat ttaatatatt gatatttata tcattttacg     2880
```

-continued

```
tttctcgttc agctttcttg tacaaagttg gcattataag aaagcattgc ttatcaattt    2940

gttgcaacga acaggtcact atcagtcaaa ataaaatcat tatttgccat ccagctgata    3000

tcccctatag tgagtcgtat tacatggtca tagctgtttc ctggcagctc tggcccgtgt    3060

ctcaaaatct ctgatgttac attgcacaag ataaaataat atcatcatga tcagtcctgc    3120

tcctcggcca cgaagtgcac gcagttgccg gccgggtcgc gcagggcgaa ctcccgcccc    3180

cacggctgct cgccgatctc ggtcatggcc ggcccggagg cgtcccggaa gttcgtggac    3240

acgacctccg accactcggc gtacagctcg tccaggccgc gcacccacac ccaggccagg    3300

gtgttgtccg gcaccacctg gtcctggacc gcgctgatga acagggtcac gtcgtcccgg    3360

accacaccgg cgaagtcgtc ctccacgaag tcccgggaga acccgagccg gtcggtccag    3420

aactcgaccg ctccggcgac gtcgcgcgcg gtgagcaccg gaacggcact ggtcaacttg    3480

gccatggttt agttcctcac cttgtcgtat tatactatgc cgatatacta tgccgatgat    3540

taattgtcaa cacgtgctga tcatgaccaa aatcccttaa cgtgagttac gcgtcgttcc    3600

actgagcgtc agaccccgta gaaaagatca aaggatcttc ttgagatcct ttttttctgc    3660

gcgtaatctg ctgcttgcaa acaaaaaaac caccgctacc agcggtggtt tgtttgccgg    3720

atcaagagct accaactctt tttccgaagg taactggctt cagcagagcg cagataccaa    3780

atactgttct tctagtgtag ccgtagttag gccaccactt caagaactct gtagcaccgc    3840

ctacatacct cgctctgcta atcctgttac cagtggctgc tgccagtggc gataagtcgt    3900

gtcttaccgg gttggactca agacgatagt taccggataa ggcgcagcgg tcgggctgaa    3960

cggggggttc gtgcacacag cccagcttgg agcgaacgac ctacaccgaa ctgagatacc    4020

tacagcgtga gctatgagaa agcgccacgc ttcccgaagg gagaaaggcg gacaggtatc    4080

cggtaagcgg cagggtcgga acaggagagc gcacgaggga gcttccaggg ggaaacgcct    4140

ggtatcttta tagtcctgtc gggtttcgcc acctctgact tgagcgtcga ttttttgtgat    4200

gctcgtcagg ggggcggagc ctatggaaaa acgccagcaa cgcggccttt ttacggttcc    4260

tggccttttg ctggcctttt gctcacatgt t                                  4291
```

```
<210> SEQ ID NO 29
<211> LENGTH: 4762
<212> TYPE: DNA
<213> ORGANISM: artificial
<220> FEATURE:
<223> OTHER INFORMATION: vector

<400> SEQUENCE: 29
```

```
ctttcctgcg ttatcccctg attctgtgga taaccgtatt accgcctttg agtgagctga    60

taccgctcgc cgcagccgaa cgaccgagcg cagcgagtca gtgagcgagg aagcggaaga    120

gcgcccaata cgcaaaccgc ctctccccgc gcgttggccg attcattaat gcagctggca    180

cgacaggttt cccgactgga aagcgggcag tgagcgcaac gcaattaata cgcgtaccgc    240

tagccaggaa gagtttgtag aaacgcaaaa aggccatccg tcaggatggc cttctgctta    300

gtttgatgcc tggcagttta tggcggggcgt cctgcccgcc accctccggg ccgttgcttc    360

acaacgttca aatcgctcc cggcggattt gtcctactca ggagagcgtt caccgacaaa    420

caacagataa aacgaaaggc ccagtcttcc gactgagcct ttcgtttttat ttgatgcctg    480

gcagttccct actctcgcgt taacgctagc atggatgttt cccagtcac gacgttgtaa    540
```

-continued

```
aacgacggcc agtcttaagc tcgggcccca aataatgatt ttattttgac tgatagtgac    600

ctgttcgttg caacacattg atgagcaatg ctttttttata atgccaactt tgtacaaaaa    660

agctgaacga gaaacgtaaa atgatataaa tatcaatata ttaaattaga ttttgcataa    720

aaaacagact acataatact gtaaaacaca acatatccag tcactatgaa tcaactactt    780

agatggtatt agtgacctgt agtcgaccga cagccttcca aatgttcttc gggtgatgct    840

gccaacttag tcgaccgaca gccttccaaa tgttcttctc aaacggaatc gtcgtatcca    900

gcctactcgc tattgtcctc aatgccgtat taaatcataa aaagaaataa gaaaaagagg    960

tgcgagcctc ttttttgtgt gacaaaataa aaacatctac ctattcatat acgctagtgt   1020

catagtcctg aaaatcatct gcatcaagaa caatttcaca actcttatac tttttctctta  1080

caagtcgttc ggcttcatct ggattttcag cctctatact tactaaacgt gataaagttt   1140

ctgtaatttc tactgtatcg acctgcagac tggctgtgta taagggagcc tgacatttat   1200

attccccaga acatcaggtt aatggcgttt ttgatgtcat tttcgcggtg gctgagatca   1260

gccacttctt ccccgataac ggagaccggc acactggcca tatcggtggt catcatgcgc   1320

cagctttcat ccccgatatg caccaccggg taaagttcac gggagacttt atctgacagc   1380

agacgtgcac tggccagggg gatcaccatc cgtcgcccgg gcgtgtcaat aatatcactc   1440

tgtacatcca caaacagacg ataacggctc tctcttttat aggtgtaaac cttaaactgc   1500

atttcaccag cccctgttct cgtcagcaaa agagccgttc atttcaataa accgggcgac   1560

ctcagccatc ccttcctgat tttccgcttt ccagcgttcg gcacgcagac gacgggcttc   1620

attctgcatg gttgtgctta ccagaccgga gatattgaca tcatatatgc cttgagcaac   1680

tgatagctgt cgctgtcaac tgtcactgta atacgctgct tcatagcata cctctttttg   1740

acatacttcg ggtatacata tcagtatata ttcttatacc gcaaaaatca gcgcgcaaat   1800

acgcatactg ttatctggct tttagtaagc cggatccacg cggcgtttac gccccgccct   1860

gccactcatc gcagtactgt tgtaattcat taagcattct gccgacatgg aagccatcac   1920

agacggcatg atgaacctga atcgccagcg gcatcagcac cttgtcgcct tgcgtataat   1980

atttgcccat ggtgaaaacg ggggcgaaga agttgtccat attggccacg tttaaatcaa   2040

aactggtgaa actcacccag ggattggctg agacgaaaaa catattctca ataaacccctt  2100

tagggaaata ggccaggttt tcaccgtaac acgccacatc ttgcgaatat atgtgtagaa   2160

actgccggaa atcgtcgtgg tattcactcc agagcgatga aaacgtttca gtttgctcat   2220

ggaaacggt gtaacaaggg tgaacactat cccatatcac cagctcaccg tctttcattg    2280

ccatacggaa ttccggatga gcattcatca ggcgggcaag aatgtgaata aaggccggat    2340

aaaacttgtg cttattttttc tttacggtct ttaaaaaggc cgtaatatcc agctgaacgg   2400

tctggttata ggtacattga gcaactgact gaaatgcctc aaaatgttct ttacgatgcc   2460

attgggatat atcaacggtg gtatatccag tgatttttttt ctccattta gcttccttag    2520

ctcctgaaaa tctcgataac tcaaaaaata cgcccggtag tgatcttatt tcattatggt   2580

gaaagttgga acctcttacg tgccgatcaa cgtctcattt cgccaaaag ttggcccagg    2640

gcttcccggt atcaacaggg acaccaggat ttatttattc tgcgaagtga tcttccgtca   2700

caggtatttta ttcggcgcaa agtgcgtcgg gtgatgctgc caacttagtc gactacaggt   2760

cactaatacc atctaagtag ttgattcata gtgactggat atgttgtgtt ttacagtatt   2820
```

-continued

```
atgtagtctg tttttttatgc aaaatctaat ttaatatatt gatatttata tcattttacg    2880

tttctcgttc agctttcttg tacaaagttg gcattataag aaagcattgc ttatcaattt    2940

gttgcaacga acaggtcact atcagtcaaa ataaaatcat tatttgccat ccagctgata    3000

tcccctatag tgagtcgtat tacatggtca tagctgtttc ctggcagctc tggcccgtgt    3060

ctcaaaatct ctgatgttac attgcacaag ataaaataat atcatcatga acaataaaac    3120

tgtctgctta cataaacagt aatacaaggg gtgttatgag ccatattcaa cgggaaacgt    3180

cgaggccgcg attaaattcc aacatggatg ctgatttata tgggtataaa tgggctcgcg    3240

ataatgtcgg gcaatcaggt gcgacaatct atcgcttgta tgggaagccc gatgcgccag    3300

agttgtttct gaaacatggc aaaggtagcg ttgccaatga tgttacagat gagatggtca    3360

gactaaactg gctgacggaa tttatgcctc ttccgaccat caagcatttt atccgtactc    3420

ctgatgatgc atggttactc accactgcga tccccggaaa aacagcattc caggtattag    3480

aagaatatcc tgattcaggt gaaaatattg ttgatgcgct ggcagtgttc ctgcgccggt    3540

tgcattcgat tcctgtttgt aattgtcctt ttaacagcga tcgcgtattt cgtctcgctc    3600

aggcgcaatc acgaatgaat aacggtttgg ttgatgcgag tgattttgat gacgagcgta    3660

atggctggcc tgttgaacaa gtctggaaag aaatgcataa acttttgcca ttctcaccgg    3720

attcagtcgt cactcatggt gatttctcac ttgataacct tatttttgac gaggggaaat    3780

taataggttg tattgatgtt ggacgagtcg gaatcgcaga ccgataccag gatcttgcca    3840

tcctatggaa ctgcctcggt gagttttctc cttcattaca gaaacggctt tttcaaaaat    3900

atggtattga taatcctgat atgaataaat tgcagtttca tttgatgctc gatgagtttt    3960

tctaatcaga attggttaat tggttgtaac actggcagag cattacgctg acttgacggg    4020

acggcgcaag ctcatgacca aaatccctta acgtgagtta cgcgtcgttc cactgagcgt    4080

cagacccccgt agaaaagatc aaaggatctt cttgagatcc ttttttttctg cgcgtaatct    4140

gctgcttgca acaaaaaaa ccaccgctac cagcggtggt ttgtttgccg gatcaagagc    4200

taccaactct ttttccgaag gtaactggct tcagcagagc gcagatacca aatactgttc    4260

ttctagtgta gccgtagtta ggccaccact tcaagaactc tgtagcaccg cctacatacc    4320

tcgctctgct aatcctgtta ccagtggctg ctgccagtgg cgataagtcg tgtcttaccg    4380

ggttggactc aagacgatag ttaccggata aggcgcagcg tcgggctga acggggggtt    4440

cgtgcacaca gcccagcttg gagcgaacga cctacaccga actgagatac ctacagcgtg    4500

agctatgaga aagcgccacg cttcccgaag ggagaaaggc ggacaggtat ccggtaagcg    4560

gcagggtcgg aacaggagag cgcacgaggg agcttccagg gggaaacgcc tggtatcttt    4620

atagtcctgt cgggtttcgc cacctctgac ttgagcgtcg atttttgtga tgctcgtcag    4680

gggggcggag cctatggaaa aacgccagca acgcggcctt tttacggttc ctggcctttt    4740

gctggccttt tgctcacatg tt    4762
```

<210> SEQ ID NO 30
<211> LENGTH: 9142
<212> TYPE: DNA
<213> ORGANISM: artificial
<220> FEATURE:
<223> OTHER INFORMATION: vector

<400> SEQUENCE: 30

-continued

```
ttatttgtct tctggttctg actctctttc tctcgtttca atgccaggtt gcctactccc      60

acaccactca caagaagatt ctactgttag tattaaatat tttttaatgt attaaatgat     120

gaatgctttt gtaaacagaa caagactatg tctaataagt gtcttgcaac atttttttaag    180

aaattaaaaa aaatatattt attatcaaaa tcaaatgtat gaaaaatcat gaataatata     240

attttataca tttttttaaa aaatctttta atttcttaat taatatctta aaaataatga     300

ttaatattta acccaaaata attagtatga ttggtaagga agatatccat gttatgtttg     360

gatgtgagtt tgatctagag caaagcttac tagagtcgac ctgcagcccc tccaccgcgg     420

tggcggccgc tctagagatc cgtcaacatg gtggagcacg acactctcgt ctactccaag    480

aatatcaaag atacagtctc agaagaccaa agggctattg agactttca acaaagggta     540

atatcgggaa acctcctcgg attccattgc ccagctatct gtcacttcat caaaaggaca     600

gtagaaaagg aaggtggcac ctacaaatgc catcattgcg ataaaggaaa ggctatcgtt     660

caagatgcct ctgccgacag tggtcccaaa gatggacccc cacccacgag gagcatcgtg     720

gaaaaagaag acgttccaac cacgtcttca aagcaagtgg attgatgtga tgatcctatg     780

cgtatggtat gacgtgtgtt caagatgatg acttcaaacc tacctatgac gtatggtatg     840

acgtgtgtcg actgatgact tagatccact cgagcggcta taaatacgta cctacgcacc     900

ctgcgctacc atccctagag ctgcagctta tttttacaac aattaccaac aacaacaaac     960

aacaaacaac attacaatta ctatttacaa ttacagtcga cccatcaaca agtttgtaca    1020

aaaaagctga acgagaaacg taaaatgata taaatatcaa tatattaaat tagattttgc    1080

ataaaaaaca gactacataa tactgtaaaa cacaacatat ccagtcatat tggcggccgc    1140

attaggcacc ccaggcttta cactttatgc ttccggctcg tataatgtgt ggattttgag    1200

ttaggatccg tcgagatttt caggagctaa ggaagctaaa atggagaaaa aaatcactgg    1260

atataccacc gttgatatat cccaatggca tcgtaaagaa catttttgagg catttcagtc    1320

agttgctcaa tgtacctata accagaccgt tcagctggat attacggcct ttttaaagac    1380

cgtaaagaaa aataagcaca gttttatcc ggcctttatt cacattcttg cccgcctgat    1440

gaatgctcat ccggaattcc gtatggcaat gaaagacggt gagctggtga tatgggatag    1500

tgttcaccct tgttacaccg tttttccatga gcaaactgaa acgttttcat cgctctggag    1560

tgaataccac gacgatttcc ggcagtttct acacatatat tcgcaagatg tggcgtgtta    1620

cggtgaaaac ctggcctatt tccctaaagg gtttattgag aatatgtttt tcgtctcagc    1680

caatccctgg gtgagtttca ccagttttga tttaaacgtg gccaatatgg acaacttctt    1740

cgcccccgtt ttcaccatgg gcaaatatta tacgcaaggc gacaaggtgc tgatgccgct    1800

ggcgattcag gttcatcatg ccgtttgtga tggcttccat gtcggcagaa tgcttaatga    1860

attacaacag tactgcgatg agtggcaggg cggggcgtaa agatctggat ccggcttact    1920

aaaagccaga taacagtatg cgtatttgcg cgctgatttt tgcggtataa gaatatatac    1980

tgatatgtat acccgaagta tgtcaaaaag aggtatgcta tgaagcagcg tattacagtg    2040

acagttgaca gcgacagcta tcagttgctc aaggcatata tgatgtcaat atctccggtc    2100

tggtaagcac aaccatgcag aatgaagccc gtcgtctgcg tgccgaacgc tggaaagcgg    2160

aaaatcagga aggatggct gaggtcgccc ggtttattga atgaacggc tcttttgctg    2220

acgagaacag gggctggtga atgcagtttt aaggtttaca cctataaaag agagagccgt    2280
```

-continued

```
tatcgtctgt ttgtggatgt acagagtgat attattgaca cgcccgggcg acggatggtg      2340

atccccctgg ccagtgcacg tctgctgtca gataaagtct cccgtgaact ttacccggtg      2400

gtgcatatcg gggatgaaag ctggcgcatg atgaccaccg atatggccag tgtgccggtc      2460

tccgttatcg gggaagaagt ggctgatctc agccaccgcg aaaatgacat caaaaacgcc      2520

attaacctga tgttctgggg aatataaatg tcaggctccc ttatacacag ccagtctgca      2580

ggtcgaccat agtgactgga tatgttgtgt tttacagtat tatgtagtct gttttttatg      2640

caaaatctaa tttaatatat tgatatttat atcattttac gtttctcgtt cagctttctt      2700

gtacaaagtg gttgataacc tagacttgtc catcttctgg attggccaac ttaattaatg      2760

tatgaaataa aaggatgcac acatagtgac atgctaatca ctataatgtg ggcatcaaag      2820

ttgtgtgtta tgtgtaatta ctagttatct gaataaaaga gaaagagatc atccatattt      2880

cttatcctaa atgaatgtca cgtgtcttta taattctttg atgaaccaga tgcatttcat      2940

taaccaaatc catatacata taaatattaa tcatatataa ttaatatcaa ttgggttagc      3000

aaaacaaatc tagtctaggt gtgttttgcg aattcgatat caagcttgat gggtaccggc      3060

gcgcccgatc atccggatat agttcctcct tcagcaaaa aacccctcaa gacccgtttta      3120

gaggccccaa ggggttatgc tagttattgc tcagcggtgg cagcagccaa ctcagcttcc      3180

tttcgggctt tgttagcagc cggatcgatc caagctgtac ctcactattc ctttgccctc      3240

ggacgagtgc tggggcgtcg gtttccacta tcggcgagta cttctacaca gccatcggtc      3300

cagacggccg cgcttctgcg ggcgatttgt gtacgcccga cagtcccggc tccggatcgg      3360

acgattgcgt cgcatcgacc ctgcgcccaa gctgcatcat cgaaattgcc gtcaaccaag      3420

ctctgataga gttggtcaag accaatgcgg agcatatacg cccggagccg cggcgatcct      3480

gcaagctccg gatgcctccg ctcgaagtag cgcgtctgct gctccataca agccaaccac      3540

ggcctccaga agaagatgtt ggcgacctcg tattgggaat ccccgaacat cgcctcgctc      3600

cagtcaatga ccgctgttat gcggccattg tccgtcagga cattgttgga gccgaaatcc      3660

gcgtgcacga ggtgccggac ttcggggcag tcctcggccc aaagcatcag ctcatcgaga      3720

gcctgcgcga cggacgcact gacggtgtcg tccatcacag tttgccagtg atacacatgg      3780

ggatcagcaa tcgcgcatat gaaatcacgc catgtagtgt attgaccgat tccttgcggt      3840

ccgaatgggc cgaacccgct cgtctggcta agatcggccg cagcgatcgc atccatagcc      3900

tccgcgaccg gctgcagaac agcgggcagt tcggtttcag gcaggtcttg caacgtgaca      3960

ccctgtgcac ggcgggagat gcaataggtc aggctctcgc tgaattcccc aatgtcaagc      4020

acttccggaa tcgggagcgc ggccgatgca aagtgccgat aaacataacg atctttgtag      4080

aaaccatcgg cgcagctatt tacccgcagg acatatccac gccctcctac atcgaagctg      4140

aaagcacgag attcttcgcc ctccgagagc tgcatcaggt cggagacgct gtcgaacttt      4200

tcgatcagaa acttctcgac agacgtcgcg gtgagttcag cttttccat gggtatatct      4260

ccttcttaaa gttaaacaaa attatttcta gagggaaacc gttgtggtct ccctatagtg      4320

agtcgtatta atttcgcggg atcgagatct gatcaacctg cattaatgaa tcggccaacg      4380

cgcggggaga ggcggtttgc gtattgggcg ctcttccgct tcctcgctca ctgactcgct      4440

gcgctcggtc gttcggctgc ggcgagcggt atcagctcac tcaaaggcgg taatacggtt      4500

atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc      4560
```

-continued

```
caggaaccgt aaaaaggccg cgttgctggc gttttttccat aggctccgcc ccccctgacga   4620

gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac tataaagata   4680

ccaggcgttt cccccctggaa gctccctcgt gcgctctcct gttccgaccc tgccgcttac   4740

cggatacctg tccgccttttc tcccttcggg aagcgtggcg ctttctcaat gctcacgctg   4800

taggtatctc agttcggtgt aggtcgttcg ctccaagctg gctgtgtgc acgaacccccc   4860

cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca acccggtaag   4920

acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag cgaggtatgt   4980

aggcggtgct acagagttct tgaagtggtg gcctaactac ggctacacta gaaggacagt   5040

atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg   5100

atccggcaaa caaaccaccg ctggtagcgg tggtttttttt gtttgcaagc agcagattac   5160

gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt ctgacgctca   5220

gtggaacgaa aactcacgtt aagggatttt ggtcatgaca ttaacctata aaaataggcg   5280

tatcacgagg ccctttcgtc tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat   5340

gcagctcccg gagacggtca cagcttgtct gtaagcggat gccgggagca gacaagcccg   5400

tcagggcgcg tcagcgggtg ttggcgggtg tcggggctgg cttaactatg cggcatcaga   5460

gcagattgta ctgagagtgc accatatgga catattgtcg ttagaacgcg gctacaatta   5520

atacataacc ttatgtatca tacacatacg atttaggtga cactatagaa cggcgcgcca   5580

agctgggtct agaactagaa acgtgatgcc acttgttatt gaagtcgatt acagcatcta   5640

ttctgtttta ctatttataa ctttgccatt tctgactttt gaaaactatc tctggatttc   5700

ggtatcgctt tgtgaagatc gagcaaaaga gacgttttgt ggacgcaatg gtccaaatcc   5760

gttctacatg aacaaattgg tcacaatttc cactaaaagt aaataaatgg caagttaaaa   5820

aaggaatatg cattttactg attgcctagg tgagctccaa gagaagttga atctacacgt   5880

ctaccaaccg ctaaaaaaag aaaaacattg aatatgtaac ctgattccat tagctttttga   5940

cttcttcaac agattctcta cttagatttc taacagaaat attattacta gcacatcatt   6000

ttcagtctca ctacagcaaa aaatccaacg gcacaataca gacaacagga gatatcagac   6060

tacagagata gatagatgct actgcatgta gtaagttaaa taaaaggaaa ataaaatgtc   6120

ttgctaccaa aactactaca gactatgatg ctcaccacag gccaaatcct gcaactagga   6180

cagcattatc ttatatatat tgtacaaaac aagcatcaag gaacatttgg tctaggcaat   6240

cagtacctcg ttctaccatc accctcagtt atcacatcct tgaaggatcc attactggga   6300

atcatcggca acacatgctc ctgatggggc acaatgacat caagaaggta ggggccaggg   6360

gtgtccaaca ttctctgaat tgccgctcta agctcttcct tcttcgtcac tcgcgctgcc   6420

ggtatcccac aagcatcagc aaacttgagc atgtttggga atatctcgct ctcgctagac   6480

ggatctccaa gataggtgtg agctctattg gacttgtaga acctatcctc caactgaacc   6540

accatacccca aatgctgatt gttcaacaac aatatcttaa ctgggagatt ctccactctt   6600

atagtggcca actcctgaac attcatgatg aaactaccat ccccatcaat gtcaaccaca   6660

acagccccag ggttagcaac agcagcacca atagccgcag gcaatccaaa acccatggct   6720

ccaagacccc ctgaggtcaa ccactgcctc ggtctcttgt acttgtaaaa ctgcgcagcc   6780

cacatttgat gctgcccaac cccagtacta acaatagcat ctccattagt caactcatca   6840
```

-continued

```
agaacctcga tagcatgctg cggagaaatc gcgtcctgga atgtcttgta acccaatgga    6900

aacttgtgtt tctgcacatt aatctcttct ctccaacctc caagatcaaa cttaccctcc    6960

actcctttct cctccaaaat catattaatt cccttcaagg ccaacttcaa atccgcgcaa    7020

accgacacgt gcgcctgctt gttcttccca atctcggcag aatcaatatc aatgtgaaca    7080

atcttagccc tactagcaaa agcctcaagc ttcccagtaa cacggtcatc aaaccttacc    7140

ccaaaggcaa gcaacaaatc actattgtca acagcatagt tagcataaac agtaccatgc    7200

atacccagca tctgaaggga atattcatca ccaataggaa aagttccaag acccattaaa    7260

gtgctagcaa cgggaatacc agtgagttca acaaagcgcc tcaattcagc actggaattc    7320

aaactgccac cgccgacgta gagaacgggc ttttgggcct ccatgatgag tctgacaatg    7380

tgttccaatt gggcctcggc ggggggcctg ggcagcctgg cgaggtaacc ggggaggtta    7440

acgggctcgt cccaattagg cacggcgagt tgctgctgaa cgtctttggg aatgtcgatg    7500

aggaccggac cggggcggcc ggaggtggcg acgaagaaag cctcggcgac gacgcggggg    7560

atgtcgtcga cgtcgaggat gaggtagttg tgcttcgtga tggatctgct cacctccacg    7620

atcggggttt cttggaaggc gtcggtgccg atcatccggc gggcgacctg gccggtgatg    7680

gcgacgactg ggacgctgtc cattaaagcg tcggcgaggc cgctcacgag gttggtggcg    7740

ccggggccgg aggtggcaat gcagacgccg gggaggccgg aggaacgcgc gtagccttcg    7800

gcggcgaaga cgccgccctg ctcgtggcgc gggagcacgt tgcggatggc ggcggagcgc    7860

gtgagcgcct ggtggatctc catcgacgca ccgccggggt acgcgaacac cgtcgtcacg    7920

ccctgcctct ccagcgcctc cacaaggatg tccgcgccct tgcgaggttc gccggaggcg    7980

aaccgtgaca cgaagggctc cgtggtcggc gcttccttgg tgaagggcgc cgccgtgggg    8040

ggtttggaga tggaacattt gattttgaga gcgtggttgg gtttggtgag ggtttgatga    8100

gagagaggga gggtggatct agtaatgcgt ttggggaagg tggggtgtga agaggaagaa    8160

gagaatcggg tggttctgga agcggtggcc gccattgtgt tgtgtggcat ggttatactt    8220

caaaaactgc acaacaagcc tagagttagt acctaaacag taaatttaca acagagagca    8280

aagacacatg caaaaatttc agccataaaa aaagttataa tagaatttaa agcaaaagtt    8340

tcattttta aacatatata caaacaaact ggatttgaag gaagggatta attcccctgc    8400

tcaaagtttg aattcctatt gtgacctata ctcgaataaa attgaagcct aaggaatgta    8460

tgagaaacaa gaaaacaaaa caaaactaca gacaaacaag tacaattaca aaattcgcta    8520

aaattctgta atcaccaaac cccatctcag tcagcacaag gcccaaggtt tattttgaaa    8580

taaaaaaaaa gtgattttat ttctcataag ctaaaagaaa gaaaggcaat tatgaaatga    8640

tttcgactag atctgaaagt caaacgcgta ttccgcagat attaaagaaa gagtagagtt    8700

tcacatggat cctagatgga cccagttgag gaaaaagcaa ggcaaagcaa accagaagtg    8760

caagatccga aattgaacca cggaatctag gatttggtag agggagaaga aaagtacctt    8820

gagaggtaga agagaagaga agagcagaga gatatatgaa cgagtgtgtc ttggtctcaa    8880

ctctgaagcg atacgagttt agaggggagc attgagttcc aatttatagg gaaaccgggt    8940

ggcagggtg agttaatgac ggaaaagccc ctaagtaacg agattggatt gtgggttaga    9000

ttcaaccgtt tgcatccgcg gcttagattg gggaagtcag agtgaatctc aaccgttgac    9060

tgagttgaaa attgaatgta gcaaccaatt gagccaaccc cagcctttgc cctttgattt    9120
```

-continued

```
tgatttgttt gttgcatact tt                                             9142
```

<210> SEQ ID NO 31
<211> LENGTH: 49911
<212> TYPE: DNA
<213> ORGANISM: artificial
<220> FEATURE:
<223> OTHER INFORMATION: vector

<400> SEQUENCE: 31

```
gtgcagcgtg acccggtcgt gcccctctct agagataatg agcattgcat gtctaagtta      60

taaaaaatta ccacatattt tttttgtcac acttgtttga agtgcagttt atctatcttt     120

atacatatat ttaaacttta ctctacgaat aatataatct atagtactac aataatatca     180

gtgttttaga gaatcatata aatgaacagt tagacatggt ctaaaggaca attgagtatt     240

ttgacaacag gactctacag ttttatcttt ttagtgtgca tgtgttctcc ttttttttttg     300

caaatagctt cacctatata atacttcatc cattttatta gtacatccat ttagggttta     360

gggttaatgg tttttataga ctaatttttt tagtacatct attttattct attttagcct     420

ctaaattaag aaaactaaaa ctctatttta gtttttttat ttaataattt agatataaaa     480

tagaataaaa taaagtgact aaaaattaaa caaataccct ttaagaaatt aaaaaaacta     540

aggaaacatt tttcttgttt cgagtagata atgccagcct gttaaacgcc gtcgacgagt     600

ctaacggaca ccaaccagcg aaccagcagc gtcgcgtcgg ccaagcgaa gcagacggca     660

cggcatctct gtcgctgcct ctggacccct ctcgagagtt ccgctccacc gttggacttg     720

ctccgctgtc ggcatccaga aattgcgtgg cggagcggca gacgtgagcc ggcacggcag     780

gcggcctcct cctcctctca cggcacggca gctacggggg attcctttcc caccgctcct     840

tcgctttccc ttcctcgccc gccgtaataa atagacaccc cctccacacc ctctttcccc     900

aacctcgtgt tgttcggagc gcacacacac acaaccagat ctcccccaaa tccacccgtc     960

ggcacctccg cttcaaggta cgccgctcgt cctcccccccc cccccctctc taccttctct    1020

agatcggcgt tccggtccat ggttagggcc cggtagttct acttctgttc atgtttgtgt    1080

tagatccgtg tttgtgttag atccgtgctg ctagcgttcg tacacggatg cgacctgtac    1140

gtcagacacg ttctgattgc taacttgcca gtgtttctct ttggggaatc ctgggatggc    1200

tctagccgtt ccgcagacgg gatcgatttc atgattttt ttgtttcgtt gcataggggtt    1260

tggtttgccc ttttcctttta tttcaatata tgccgtgcac ttgtttgtcg ggtcatcttt    1320

tcatgctttt ttttgtcttg gttgtgatga tgtggtctgg ttgggcggtc gttctagatc    1380

ggagtagaat tctgtttcaa actacctggt ggatttatta attttggatc tgtatgtgtg    1440

tgccatacat attcatagtt acgaattgaa gatgatggat ggaaatatcg atctaggata    1500

ggtatacatg ttgatgcggg ttttactgat gcatatacag agatgctttt tgttcgcttg    1560

gttgtgatga tgtggtgtgg ttgggcggtc gttcattcgt tctagatcgg agtagaatac    1620

tgtttcaaac tacctggtgt atttattaat tttggaactg tatgtgtgtg tcatacatct    1680

tcatagttac gagtttaaga tggatggaaa tatcgatcta ggataggtat acatgttgat    1740

gtgggtttta ctgatgcata tacatgatgg catatgcagc atctattcat atgctctaac    1800

cttgagtacc tatctattat aataaacaag tatgttttat aattattttg atcttgatat    1860

acttggatga tggcatatgc agcagctata tgtggatttt tttagccctg ccttcatacg    1920
```

-continued

```
ctatttattt gcttggtact gtttcttttg tcgatgctca ccctgttgtt tggtgttact    1980

tctgcaggtc gactctagag gatccacaag tttgtacaaa aaagctgaac gagaaacgta    2040

aaatgatata aatatcaata tattaaatta gattttgcat aaaaaacaga ctacataata    2100

ctgtaaaaca caacatatcc agtcactatg gcggccgcat taggcacccc aggctttaca    2160

ctttatgctt ccggctcgta taatgtgtgg attttgagtt aggatttaaa tacgcgttga    2220

tccggcttac taaaagccag ataacagtat gcgtatttgc gcgctgattt ttgcggtata    2280

agaatatata ctgatatgta tacccgaagt atgtcaaaaa gaggtatgct atgaagcagc    2340

gtattacagt gacagttgac agcgacagct atcagttgct caaggcatat atgatgtcaa    2400

tatctccggt ctggtaagca caaccatgca gaatgaagcc cgtcgtctgc gtgccgaacg    2460

ctggaaagcg gaaaatcagg aagggatggc tgaggtcgcc cggtttattg aaatgaacgg    2520

ctcttttgct gacgagaaca ggggctggtg aaatgcagtt taaggtttac acctataaaa    2580

gagagagccg ttatcgtctg tttgtggatg tacagagtga tatcattgac acgcccgggtc    2640

gacggatggt gatcccccctg gccagtgcac gtctgctgtc agataaagtc tcccgtgaac    2700

tttacccggt ggtgcatatc ggggatgaaa gctggcgcat gatgaccacc gatatggcca    2760

gtgtgccggt ctccgttatc ggggaagaag tggctgatct cagccaccgc gaaaatgaca    2820

tcaaaaacgc cattaacctg atgttctggg aatataaat gtcaggctcc cttatacaca    2880

gccagtctgc aggtcgacca tagtgactgg atatgttgtg ttttacagta ttatgtagtc    2940

tgttttttat gcaaaatcta atttaatata ttgatattta tatcatttta cgtttctcgt    3000

tcagctttct tgtacaaagt ggtgttaacc tagacttgtc catcttctgg attggccaac    3060

ttaattaatg tatgaaataa aaggatgcac acatagtgac atgctaatca ctataatgtg    3120

ggcatcaaag ttgtgtgtta tgtgtaatta ctagttatct gaataaaaga gaaagagatc    3180

atccatattt cttatcctaa atgaatgtca cgtgtcttta taattctttg atgaaccaga    3240

tgcatttcat taaccaaatc catatacata taaatattaa tcatatataa ttaatatcaa    3300

ttgggttagc aaaacaaatc tagtctaggt gtgttttgcg aattgcggcc gccaccgcgg    3360

tggagctcga attccggtcc gggtcacctt gtccaccaa gatggaactg cggccgctca    3420

ttaattaagt caggcgcgcc tctagttgaa gacacgttca tgtcttcatc gtaagaagac    3480

actcagtagt cttcggccag aatggccatc tggattcagc aggcctagaa ggccatttaa    3540

atcctgagga tctggtcttc ctaaggaccc gggatatcgg accgattaaa ctttaattcg    3600

gtccgaagct tgcatgcctg cagtgcagcg tgacccggtc gtgcccctct ctagagataa    3660

tgagcattgc atgtctaagt tataaaaaat taccacatat tttttttgtc acacttgttt    3720

gaagtgcagt ttatctatct ttatacatat atttaaactt tactctacga ataatataat    3780

ctatagtact acaataatat cagtgtttta gagaatcata taaatgaaca gttagacatg    3840

gtctaaagga caattgagta ttttgacaac aggactctac agttttatct tttagtgtg    3900

catgtgttct cctttttttt tgcaaatagc ttcacctata taatacttca tccattttat    3960

tagtacatcc atttagggtt tagggttaat ggttttttata gactaatttt tttagtacat    4020

ctatttatt ctattttagc ctctaaatta agaaaactaa aactctattt tagttttttt    4080

atttaataat ttagatataa aatagaataa aataaagtga ctaaaaatta aacaaatacc    4140

ctttaagaaa ttaaaaaaac taaggaaaca tttttcttgt ttcgagtaga taatgccagc    4200
```

-continued

```
ctgttaaacg ccgtcgacga gtctaacgga caccaaccag cgaaccagca gcgtcgcgtc    4260

gggccaagcg aagcagacgg cacggcatct ctgtcgctgc ctctggaccc ctctcgagag    4320

ttccgctcca ccgttggact tgctccgctg tcggcatcca gaaattgcgt ggcggagcgg    4380

cagacgtgag ccggcacggc aggcggcctc ctcctcctct cacggcaccg gcagctacgg    4440

gggattcctt tcccaccgct ccttcgcttt cccttcctcg cccgccgtaa taaatagaca    4500

ccccctccac accctctttc cccaacctcg tgttgttcgg agcgcacaca cacacaacca    4560

gatctccccc aaatccaccc gtcggcacct ccgcttcaag gtacgccgct cgtcctcccc    4620

ccccccctc tctaccttct ctagatcggc gttccggtcc atgcatggtt agggcccggt    4680

agttctactt ctgttcatgt ttgtgttaga tccgtgtttg tgttagatcc gtgctgctag    4740

cgttcgtaca cggatgcgac ctgtacgtca gacacgttct gattgctaac ttgccagtgt    4800

ttctctttgg ggaatcctgg gatggctcta gccgttccgc agacgggatc gatttcatga    4860

ttttttttgt ttcgttgcat agggtttggt ttgccctttt cctttatttc aatatatgcc    4920

gtgcacttgt ttgtcgggtc atcttttcat gcttttttttt gtcttggttg tgatgatgtg    4980

gtctggttgg gcggtcgttc tagatcggag tagaattctg tttcaaacta cctggtggat    5040

ttattaattt tggatctgta tgtgtgtgcc atacatattc atagttacga attgaagatg    5100

atggatggaa atatcgatct aggataggta tacatgttga tgcgggtttt actgatgcat    5160

atacagagat gctttttgtt cgcttggttg tgatgatgtg gtgtggttgg gcggtcgttc    5220

attcgttcta gatcggagta gaatactgtt tcaaactacc tggtgtattt attaattttg    5280

gaactgtatg tgtgtgtcat acatcttcat agttacgagt ttaagatgga tggaaatatc    5340

gatctaggat aggtatacat gttgatgtgg gttttactga tgcatataca tgatggcata    5400

tgcagcatct attcatatgc tctaaccttg agtacctatc tattataata aacaagtatg    5460

ttttataatt attttgatct tgatatactt ggatgatggc atatgcagca gctatatgtg    5520

gattttttta gccctgcctt catacgctat ttatttgctt ggtactgttt cttttgtcga    5580

tgctcaccct gttgtttggt gttacttctg caggtcgact ttaacttagc ctaggatcca    5640

cacgacacca tgtcccccga gcgccgcccc gtcgagatcc gcccggccac cgccgccgac    5700

atggccgccg tgtgcgacat cgtgaaccac tacatcgaga cctccaccgt gaacttccgc    5760

accgagccgc agaccccgca ggagtggatc gacgacctgg agcgcctcca ggaccgctac    5820

ccgtggctcg tggccgaggt ggagggcgtg gtggccggca tcgcctacgc cggcccgtgg    5880

aaggcccgca acgcctacga ctggaccgtg agtccaccg tgtacgtgtc ccaccgccac    5940

cagcgcctcg gcctcggctc caccctctac acccacctcc tcaagagcat ggaggcccag    6000

ggcttcaagt ccgtggtggc cgtgatcggc ctcccgaacg accgtccgt gcgcctccac    6060

gaggccctcg gctacaccgc ccgcggcacc ctccgcgccg ccggctacaa gcacggcggc    6120

tggcacgacg tcggcttctg gcagcgcgac ttcgagctgc cggccccgcc gcgcccggtg    6180

cgccggtga cgcagatctg agtcgaaacc tagacttgtc catcttctgg attggccaac    6240

ttaattaatg tatgaaataa aaggatgcac acatagtgac atgctaatca ctataatgtg    6300

ggcatcaaag ttgtgtgtta tgtgtaatta ctagttatct gaataaaaga gaaagagatc    6360

atccatattt cttatcctaa atgaatgtca cgtgtcttta taattctttg atgaaccaga    6420

tgcatttcat taaccaaatc catatacata taaatattaa tcatatataa ttaatatcaa    6480
```

-continued

```
ttgggttagc aaaacaaatc tagtctaggt gtgttttgcg aattgcggcc gccaccgcgg    6540

tggagctcga attcattccg attaatcgtg gcctcttgct cttcaggatg aagagctatg    6600

tttaaacgtg caagcgctac tagacaattc agtacattaa aaacgtccgc aatgtgttat    6660

taagttgtct aagcgtcaat ttggtttaca ccacaatata tcctgccacc agccagccaa    6720

cagctccccg accggcagct cggcacaaaa tcaccactcg atacaggcag cccatcagtc    6780

cgggacggcg tcagcgggag agccgttgta aggcggcaga ctttgctcat gttaccgatg    6840

ctattcggaa gaacggcaac taagctgccg ggtttgaaac acggatgatc tcgcggaggg    6900

tagcatgttg attgtaacga tgacagagcg ttgctgcctg tgatcaaata tcatctccct    6960

cgcagagatc cgaattatca gccttcttat tcatttctcg cttaaccgtg acaggctgtc    7020

gatcttgaga actatgccga cataatagga aatcgctgga taaagccgct gaggaagctg    7080

agtggcgcta tttctttaga agtgaacgtt gacgatcgtc gaccgtaccc cgatgaatta    7140

attcggacgt acgttctgaa cacagctgga tacttacttg ggcgattgtc atacatgaca    7200

tcaacaatgt acccgtttgt gtaaccgtct cttggaggtt cgtatgacac tagtggttcc    7260

cctcagcttg cgactagatg ttgaggccta acatttttatt agagagcagg ctagttgctt    7320

agatacatga tcttcaggcc gttatctgtc agggcaagcg aaaattggcc atttatgacg    7380

accaatgccc cgcagaagct cccatctttg ccgccataga cgccgcgccc ccctttttggg    7440

gtgtagaaca tccttttgcc agatgtggaa aagaagttcg ttgtcccatt gttggcaatg    7500

acgtagtagc cggcgaaagt gcgagaccca tttgcgctat atataagcct acgatttccg    7560

ttgcgactat tgtcgtaatt ggatgaacta ttatcgtagt tgctctcaga gttgtcgtaa    7620

tttgatggac tattgtcgta attgcttatg gagttgtcgt agttgcttgg agaaatgtcg    7680

tagttggatg gggagtagtc atagggaaga cgagcttcat ccactaaaac aattggcagg    7740

tcagcaagtg cctgccccga tgccatcgca agtacgaggc ttagaaccac cttcaacaga    7800

tcgcgcatag tcttccccag ctctctaacg cttgagttaa gccgcgccgc gaagcggcgt    7860

cggcttgaac gaattgttag acattatttg ccgactacct tggtgatctc gcctttcacg    7920

tagtgaacaa attcttccaa ctgatctgcg cgcgaggcca agcgatcttc ttgtccaaga    7980

taagcctgcc tagcttcaag tatgacgggc tgatactggg ccggcaggcg ctccattgcc    8040

cagtcggcag cgacatcctt cggcgcgatt ttgccggtta ctgcgctgta ccaaatgcgg    8100

gacaacgtaa gcactacatt tcgctcatcg ccagcccagt cgggcggcga gttccatagc    8160

gttaaggttt catttagcgc ctcaaataga tcctgttcag gaaccggatc aaagagttcc    8220

tccgccgctg gacctaccaa ggcaacgcta tgttctcttg cttttgtcag caagatagcc    8280

agatcaatgt cgatcgtggc tggctcgaag atacctgcaa gaatgtcatt gcgctgccat    8340

tctccaaatt gcagttcgcg cttagctgga taacgccacg gaatgatgtc gtcgtgcaca    8400

acaatggtga cttctacagc gcggagaatc tcgctctctc caggggaagc cgaagtttcc    8460

aaaaggtcgt tgatcaaagc tcgccgcgtt gtttcatcaa gccttacagt caccgtaacc    8520

agcaaatcaa tatcactgtg tggcttcagg ccgccatcca ctgcggagcc gtacaaatgt    8580

acggccagca acgtcggttc gagatggcgc tcgatgacgc caactacctc tgatagttga    8640

gtcgatactt cggcgatcac cgcttccctc atgatgttta actcctgaat taagccgcgc    8700

cgcgaagcgg tgtcggcttg aatgaattgt taggcgtcat cctgtgctcc cgagaaccag    8760
```

```
taccagtaca tcgctgtttc gttcgagact tgaggtctag tttttatacgt gaacaggtca    8820

atgccgccga gagtaaagcc acattttgcg tacaaattgc aggcaggtac attgttcgtt    8880

tgtgtctcta atcgtatgcc aaggagctgt ctgcttagtg cccacttttt cgcaaattcg    8940

atgagactgt gcgcgactcc tttgcctcgg tgcgtgtgcg acacaacaat gtgttcgata    9000

gaggctagat cgttccatgt tgagttgagt tcaatcttcc cgacaagctc ttggtcgatg    9060

aatgcgccat agcaagcaga gtcttcatca gagtcatcat ccgagatgta atccttccgg    9120

taggggctca cacttctggt agatagttca aagccttggt cggataggtg cacatcgaac    9180

acttcacgaa caatgaaatg gttctcagca tccaatgttt ccgccacctg ctcagggatc    9240

accgaaatct tcatatgacg cctaacgcct ggcacagcgg atcgcaaacc tggcgcggct    9300

tttggcacaa aaggcgtgac aggtttgcga atccgttgct gccacttgtt aacccttttg    9360

ccagatttgg taactataat ttatgttaga ggcgaagtct tgggtaaaaa ctggcctaaa    9420

attgctgggg atttcaggaa agtaaacatc accttccggc tcgatgtcta ttgtagatat    9480

atgtagtgta tctacttgat cgggggatct gctgcctcgc gcgtttcggt gatgacggtg    9540

aaaacctctg acacatgcag ctcccggaga cggtcacagc ttgtctgtaa gcggatgccg    9600

ggagcagaca agcccgtcag ggcgcgtcag cgggtgttgg cgggtgtcgg ggcgcagcca    9660

tgacccagtc acgtagcgat agcggagtgt atactggctt aactatgcgg catcagagca    9720

gattgtactg agagtgcacc atatgcggtg tgaaataccg cacagatgcg taaggagaaa    9780

ataccgcatc aggcgctctt ccgcttcctc gctcactgac tcgctgcgct cggtcgttcg    9840

gctgcggcga gcggtatcag ctcactcaaa ggcggtaata cggttatcca cagaatcagg    9900

ggataacgca ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa    9960

ggccgcgttg ctggcgtttt tccataggct ccgcccccct gacgagcatc acaaaaatcg   10020

acgctcaagt cagaggtggc gaaacccgac aggactataa agataccagg cgtttccccc   10080

tggaagctcc ctcgtgcgct ctcctgttcc gaccctgccg cttaccggat acctgtccgc   10140

ctttctccct tcgggaagcg tggcgctttc tcatagctca cgctgtaggt atctcagttc   10200

ggtgtaggtc gttcgctcca agctgggctg tgtgcacgaa ccccccgttc agcccgaccg   10260

ctgcgcctta tccggtaact atcgtcttga gtccaacccg gtaagacacg acttatcgcc   10320

actggcagca gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga   10380

gttcttgaag tggtggccta actacggcta cactagaagg acagtatttg gtatctgcgc   10440

tctgctgaag ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaacaaac   10500

caccgctggt agcggtggtt tttttgtttg caagcagcag attacgcgca gaaaaaaagg   10560

atctcaagaa gatcctttga tcttttctac ggggtctgac gctcagtgga acgaaaactc   10620

acgttaaggg attttggtca tgagattatc aaaaaggatc ttcacctaga tccttttaaa   10680

ttaaaaatga agttttaaat caatctaaag tatatatgag taaacttggt ctgacagtta   10740

ccaatgctta atcagtgagg cacctatctc agcgatctgt ctatttcgtt catccatagt   10800

tgcctgactc cccgtcgtgt agataactac gatacgggag ggcttaccat ctggccccag   10860

tgctgcaatg ataccgcgag acccacgctc accggctcca gatttatcag caataaacca   10920

gccagccgga agggccgagc gcagaagtgg tcctgcaact ttatccgcct ccatccagtc   10980

tattaattgt tgccgggaag ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt   11040
```

-continued

```
tgttgccatt gctgcagggg ggggggggg gggggacttc cattgttcat tccacggaca    11100

aaaacagaga aaggaaacga cagaggccaa aaagcctcgc tttcagcacc tgtcgtttcc    11160

tttcttttca gagggtattt taaataaaaa cattaagtta tgacgaagaa gaacggaaac    11220

gccttaaacc ggaaaatttt cataaatagc gaaaacccgc gaggtcgccg ccccgtaacc    11280

tacctgtcgg atcaccggaa aggacccgta aagtgataat gattatcatc tacatatcac    11340

aacgtgcgtg gaggccatca aaccacgtca aataatcaat tatgacgcag gtatcgtatt    11400

aattgatctg catcaactta acgtaaaaac aacttcagac aatacaaatc agcgacactg    11460

aatacggggc aacctcatgt cccccccccc cccccccctg caggcatcgt ggtgtcacgc    11520

tcgtcgtttg gtatggcttc attcagctcc ggttcccaac gatcaaggcg agttacatga    11580

tcccccatgt tgtgcaaaaa agcggttagc tccttcggtc ctccgatcgt tgtcagaagt    11640

aagttggccg cagtgttatc actcatggtt atggcagcac tgcataattc tcttactgtc    11700

atgccatccg taagatgctt ttctgtgact ggtgagtact caaccaagtc attctgagaa    11760

tagtgtatgc ggcgaccgag ttgctcttgc ccggcgtcaa cacgggataa taccgcgcca    11820

catagcagaa ctttaaaagt gctcatcatt ggaaaacgtt cttcggggcg aaaactctca    11880

aggatcttac cgctgttgag atccagttcg atgtaaccca ctcgtgcacc caactgatct    11940

tcagcatctt ttactttcac cagcgtttct gggtgagcaa aaacaggaag gcaaaatgcc    12000

gcaaaaaagg gaataagggc gacacggaaa tgttgaatac tcatactctt ccttttttcaa    12060

tattattgaa gcatttatca gggttattgt ctcatgagcg gatacatatt tgaatgtatt    12120

tagaaaaata aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc acctgacgtc    12180

taagaaacca ttattatcat gacattaacc tataaaaata ggcgtatcac gaggccctt    12240

cgtcttcaag aattcggagc ttttgccatt ctcaccggat tcagtcgtca ctcatggtga    12300

tttctcactt gataacctta ttttttgacga ggggaaatta ataggttgta ttgatgttgg    12360

acgagtcgga atcgcagacc gataccagga tcttgccatc ctatggaact gcctcggtga    12420

gttttctcct tcattacaga aacggctttt tcaaaaatat ggtattgata atcctgatat    12480

gaataaattg cagtttcatt tgatgctcga tgagttttc taatcagaat tggttaattg    12540

gttgtaacac tggcagagca ttacgctgac ttgacgggac ggcggctttg ttgaataaat    12600

cgaactttttg ctgagttgaa ggatcagatc acgcatcttc ccgacaacgc agaccgttcc    12660

gtggcaaagc aaaagttcaa aatcaccaac tggtccacct acaacaaagc tctcatcaac    12720

cgtggctccc tcactttctg gctggatgat ggggcgattc aggcctggta tgagtcagca    12780

acaccttctt cacgaggcag acctcagcgc cagaaggccg ccagagaggc gagcgcggc    12840

cgtgaggctt ggacgctagg gcagggcatg aaaaagcccg tagcgggctg ctacgggcgt    12900

ctgacgcggt ggaaaggggg aggggatgtt gtctacatgg ctctgctgta gtgagtgggt    12960

tgcgctccgg cagcggtcct gatcaatcgt caccctttct cggtccttca acgttcctga    13020

caacgagcct ccttttcgcc aatccatcga caatcaccgc gagtccctgc tcgaacgctg    13080

cgtccggacc ggcttcgtcg aaggcgtcta tcgcggcccg caacagcggc gagagcggag    13140

cctgttcaac ggtgccgccg cgctcgccgg catcgctgtc gccggcctgc tcctcaagca    13200

cggccccaac agtgaagtag ctgattgtca tcagcgcatt gacggcgtcc ccggccgaaa    13260

aacccgcctc gcagaggaag cgaagctgcg cgtcggccgt ttccatctgc ggtgcgcccg    13320
```

-continued

```
gtcgcgtgcc ggcatggatg cgcgcgccat cgcggtaggc gagcagcgcc tgcctgaagc   13380

tgcgggcatt cccgatcaga aatgagcgcc agtcgtcgtc ggctctcggc accgaatgcg   13440

tatgattctc cgccagcatg gcttcggcca gtgcgtcgag cagcgcccgc ttgttcctga   13500

agtgccagta aagcgccggc tgctgaaccc ccaaccgttc cgccagtttg cgtgtcgtca   13560

gaccgtctac gccgacctcg ttcaacaggt ccagggcggc acggatcact gtattcggct   13620

gcaactttgt catgcttgac actttatcac tgataaacat aatatgtcca ccaacttatc   13680

agtgataaag aatccgcgcg ttcaatcgga ccagcggagg ctggtccgga ggccagacgt   13740

gaaacccaac atacccctga tcgtaattct gagcactgtc gcgctcgacg ctgtcggcat   13800

cggcctgatt atgccggtgc tgccgggcct cctgcgcgat ctggttcact cgaacgacgt   13860

caccgcccac tatggcattc tgctggcgct gtatgcgttg gtgcaatttg cctgcgcacc   13920

tgtgctgggc gcgctgtcgg atcgtttcgg cggcggcca atcttgctcg tctcgctggc   13980

cggcgccact gtcgactacg ccatcatggc gacagcgcct ttcctttggg ttctctatat   14040

cgggcggatc gtggccggca tcaccggggc gactggggcg gtagccggcg cttatattgc   14100

cgatatcact gatggcgatg agcgcgcgcg gcacttcggc ttcatgagcg cctgtttcgg   14160

gttcgggatg gtcgcgggac ctgtgctcgg tgggctgatg ggcggtttct ccccccacgc   14220

tccgttcttc gccgcggcag ccttgaacgg cctcaatttc ctgacgggct gtttcctttt   14280

gccggagtcg cacaaaggcg aacgccggcc gttacgccgg gaggctctca acccgctcgc   14340

ttcgttccgg tgggcccggg gcatgaccgt cgtcgccgcc ctgatggcgg tcttcttcat   14400

catgcaactt gtcggacagg tgccggccgc gctttgggtc attttcggcg aggatcgctt   14460

tcactgggac gcgaccacga tcggcatttc gcttgccgca tttggcattc tgcattcact   14520

cgcccaggca atgatcaccg ccctgtagc cgcccggctc ggcgaaaggc gggcactcat   14580

gctcggaatg attgccgacg gcacaggcta catcctgctt gccttcgcga cacggggatg   14640

gatggcgttc ccgatcatgg tcctgcttgc ttcgggtggc atcggaatgc cggcgctgca   14700

agcaatgttg tccaggcagg tggatgagga acgtcagggg cagctgcaag gctcactggc   14760

ggcgctcacc agcctgacct cgatcgtcgg acccctcctc ttcacggcga tctatgcggc   14820

ttctataaca acgtggaacg ggtgggcatg gattgcaggc gctgccctct acttgctctg   14880

cctgccggcg ctgcgtcgcg ggctttggag cggcgcaggg caacgagccg atcgctgatc   14940

gtggaaacga taggcctatg ccatgcgggt caaggcgact tccggcaagc tatacgcgcc   15000

ctaggagtgc ggttggaacg ttggcccagc cagatactcc cgatcacgag caggacgccg   15060

atgatttgaa gcgcactcag cgtctgatcc aagaacaacc atcctagcaa cacggcggtc   15120

cccgggctga aaagcccag taaggaaaca actgtaggtt cgagtcgcga atcccccgg   15180

aaccaaagga agtaggttaa acccgctccg atcaggccga ccacgccag gccgagaaca   15240

ttggttcctg taggcatcgg gattggcgga tcaaacacta aagctactgg aacgagcaga   15300

agtcctccgg ccgccagttg ccaggcggta aaggtgagca gaggcacggg aggttgccac   15360

ttgcgggtca gcacggttcc gaacgccatg gaaaccgccc ccgccaggcc cgctgcgacg   15420

ccgacaggat ctagcgctgc gtttggtgtc aacaccaaca gcgccacgcc cgcagttccg   15480

caaatagccc ccaggaccgc catcaatcgt atcgggctac ctagcagagc ggcagagatg   15540

aacacgacca tcagcggctg cacagcgcct accgtcgccg cgaccccgcc cggcaggcgg   15600
```

-continued

```
tagaccgaaa taaacaacaa gctccagaat agcgaaatat taagtgcgcc gaggatgaag    15660

atgcgcatcc accagattcc cgttggaatc tgtcggacga tcatcacgag caataaaccc    15720

gccggcaacg cccgcagcag cataccggcg acccctcggc ctcgctgttc gggctccacg    15780

aaaacgccgg acagatgcgc cttgtgagcg tccttggggc cgtcctcctg tttgaagacc    15840

gacagcccaa tgatctcgcc gtcgatgtag gcgccgaatg ccacggcatc tcgcaaccgt    15900

tcagcgaacg cctccatggg cttttctcc tcgtgctcgt aaacggaccc gaacatctct     15960

ggagctttct tcagggccga caatcggatc tcgcggaaat cctgcacgtc ggccgctcca    16020

agccgtcgaa tctgagcctt aatcacaatt gtcaatttta atcctctgtt tatcggcagt    16080

tcgtagagcg cgccgtgcgt cccgagcgat actgagcgaa gcaagtgcgt cgagcagtgc    16140

ccgcttgttc ctgaaatgcc agtaaagcgc tggctgctga accccagcc ggaactgacc      16200

ccacaaggcc ctagcgtttg caatgcacca ggtcatcatt gacccaggcg tgttccacca    16260

ggccgctgcc tcgcaactct tcgcaggctt cgccgacctg ctcgcgccac ttcttcacgc    16320

gggtggaatc cgatccgcac atgaggcgga aggtttccag cttgagcggg tacggctccc    16380

ggtgcgagct gaaatagtcg aacatccgtc gggccgtcgg cgacagcttg cggtacttct    16440

cccatatgaa tttcgtgtag tggtcgccag caaacagcac gacgatttcc tcgtcgatca    16500

ggacctggca acgggacgtt ttcttgccac ggtccaggac gcggaagcgg tgcagcagcg    16560

acaccgattc caggtgccca acgcggtcgg acgtgaagcc catcgccgtc gcctgtaggc    16620

gcgacaggca ttcctcggcc ttcgtgtaat accggccatt gatcgaccag cccaggtcct    16680

ggcaaagctc gtagaacgtg aaggtgatcg gctcgccgat aggggtgcgc ttcgcgtact    16740

ccaacacctg ctgccacacc agttcgtcat cgtcggcccg cagctcgacg ccggtgtagg    16800

tgatcttcac gtccttgttg acgtggaaaa tgaccttgtt ttgcagcgcc tcgcgcggga    16860

ttttcttgtt gcgcgtggtg aacagggcag agcgggccgt gtcgtttggc atcgctcgca    16920

tcgtgtccgg ccacggcgca atatcgaaca aggaaagctg catttccttg atctgctgct    16980

tcgtgtgttt cagcaacgcg gcctgcttgg cctcgctgac ctgttttgcc aggtcctcgc    17040

cggcggtttt tcgcttcttg gtcgtcatag ttcctcgcgt gtcgatggtc atcgacttcg    17100

ccaaacctgc cgcctcctgt tcgagacgac gcgaacgctc cacggcggcc gatggcgcgg    17160

gcagggcagg gggagccagt tgcacgctgt cgcgctcgat cttggccgta gcttgctgga    17220

ccatcgagcc gacggactgg aaggtttcgc ggggcgcacg catgacggtg cggcttgcga    17280

tggtttcggc atcctcggcg gaaaaccccg cgtcgatcag ttcttgcctg tatgccttcc    17340

ggtcaaacgt ccgattcatt caccctcctt gcgggattgc cccgactcac gccgggggcaa   17400

tgtgcccta ttcctgattt gacccgcctg gtgccttggt gtccagataa tccaccttat      17460

cggcaatgaa gtcggtcccg tagaccgtct ggccgtcctt ctcgtacttg gtattccgaa    17520

tcttgccctg cacgaatacc agcgacccct tgcccaaata cttgccgtgg gcctcggcct    17580

gagagccaaa acacttgatg cggaagaagt cggtgcgctc ctgcttgtcg ccggcatcgt    17640

tgcgccactc ttcattaacc gctatatcga aaattgcttg cggcttgtta gaattgccat    17700

gacgtacctc ggtgtcacgg gtaagattac cgataaactg gaactgatta tggctcatat    17760

cgaaagtctc cttgagaaag gagactctag tttagctaaa cattggttcc gctgtcaaga    17820

actttagcgg ctaaaatttt gcgggccgcg accaaaggtg cgaggggcgg cttccgctgt    17880
```

-continued

```
gtacaaccag atatttttca ccaacatcct tcgtctgctc gatgagcggg gcatgacgaa   17940

acatgagctg tcggagaggg caggggtttc aatttcgttt ttatcagact taaccaacgg   18000

taaggccaac ccctcgttga aggtgatgga ggccattgcc gacgccctgg aaactcccct   18060

acctcttctc ctggagtcca ccgaccttga ccgcgaggca ctcgcggaga ttgcgggtca   18120

tcctttcaag agcagcgtgc cgcccggata cgaacgcatc agtgtggttt tgccgtcaca   18180

taaggcgttt atcgtaaaga aatggggcga cgacacccga aaaaagctgc gtggaaggct   18240

ctgacgccaa gggttagggc ttgcacttcc ttctttagcc gctaaaacgg cccctttctct   18300

gcgggccgtc ggctcgcgca tcatatcgac atcctcaacg gaagccgtgc cgcgaatggc   18360

atcgggcggg tgcgctttga cagttgtttt ctatcagaac ccctacgtcg tgcggttcga   18420

ttagctgttt gtcttgcagg ctaaacactt tcggtatatc gtttgcctgt gcgataatgt   18480

tgctaatgat ttgttgcgta ggggttactg aaaagtgagc gggaaagaag agtttcagac   18540

catcaaggag cgggccaagc gcaagctgga acgcgacatg ggtgcggacc tgttggccgc   18600

gctcaacgac ccgaaaaccg ttgaagtcat gctcaacgcg gacggcaagg tgtggcacga   18660

acgccttggc gagccgatgc ggtacatctg cgacatgcgg cccagccagt cgcaggcgat   18720

tatagaaacg gtggccggat tccacggcaa agaggtcacg cggcattcgc ccatcctgga   18780

aggcgagttc cccttggatg gcagccgctt tgccggccaa ttgccgccgg tcgtggccgc   18840

gccaaccttt gcgatccgca agcgcgcggt cgccatcttc acgctggaac agtacgtcga   18900

ggcgggcatc atgacccgcg agcaatacga ggtcattaaa agcgccgtcg cggcgcatcg   18960

aaacatcctc gtcattggcg gtactggctc gggcaagacc acgctcgtca acgcgatcat   19020

caatgaaatg gtcgccttca acccgtctga gcgcgtcgtc atcatcgagg acaccggcga   19080

aatccagtgc gccgcagaga acgccgtcca ataccacacc agcatcgacg tctcgatgac   19140

gctgctgctc aagacaacgc tgcgtatgcg ccccgaccgc atcctggtcg gtgaggtacg   19200

tggccccgaa gcccttgatc tgttgatggc ctggaacacc gggcatgaag gaggtgccgc   19260

caccctgcac gcaaacaacc ccaaagcggg cctgagccgg ctcgccatgc ttatcagcat   19320

gcacccggat tcaccgaaac ccattgagcc gctgattggc gaggcggttc atgtggtcgt   19380

ccatatcgcc aggaccccta gcggccgtcg agtgcaagaa attctcgaag ttcttggtta   19440

cgagaacggc cagtacatca ccaaaaccct gtaaggagta tttccaatga caacggctgt   19500

tccgttccgt ctgaccatga atcgcggcat tttgttctac cttgccgtgt tcttcgttct   19560

cgctctcgcg ttatccgcgc atccggcgat ggcctcggaa ggcaccggcg gcagcttgcc   19620

atatgagagc tggctgacga acctgcgcaa ctccgtaacc ggcccggtgg ccttcgcgct   19680

gtccatcatc ggcatcgtcg tcgcggcggg cgtgctgatc ttcggcggcg aactcaacgc   19740

cttcttccga accctgatct tcctggttct ggtgatggcg ctgctggtcg gcgcgcagaa   19800

cgtgatgagc accttcttcg gtcgtggtgc cgaaatcgcg gccctcggca acggggcgct   19860

gcaccaggtg caagtcgcgg cggcggatgc cgtgcgtgcg gtagcggctg gacggctcgc   19920

ctaatcatgg ctctgcgcac gatccccatc cgtcgcgcag gcaaccgaga aaacctgttc   19980

atgggtggtg atcgtgaact ggtgatgttc tcgggcctga tggcgtttgc gctgattttc   20040

agcgcccaag agctgcgggc caccgtggtc ggtctgatcc tgtggttcgg ggcgctctat   20100

gcgttccgaa tcatggcgaa ggccgatccg aagatgcggt tcgtgtacct gcgtcaccgc   20160
```

-continued

```
cggtacaagc cgtattaccc ggcccgctcg accccgttcc gcgagaacac caatagccaa   20220

gggaagcaat accgatgatc caagcaattg cgattgcaat cgcgggcctc ggcgcgcttc   20280

tgttgttcat cctctttgcc cgcatccgcg cggtcgatgc cgaactgaaa ctgaaaaagc   20340

atcgttccaa ggacgccggc ctggccgatc tgctcaacta cgccgctgtc gtcgatgacg   20400

gcgtaatcgt gggcaagaac ggcagcttta tggctgcctg gctgtacaag ggcgatgaca   20460

acgcaagcag caccgaccag cagcgcgaag tagtgtccgc ccgcatcaac caggccctcg   20520

cgggcctggg aagtgggtgg atgatccatg tggacgccgt gcggcgtcct gctccgaact   20580

acgcggagcg gggcctgtcg gcgttccctg accgtctgac ggcagcgatt gaagaagagc   20640

gctcggtctt gccttgctcg tcggtgatgt acttcaccag ctccgcgaag tcgctcttct   20700

tgatggagcg catggggacg tgcttggcaa tcacgcgcac cccccggccg ttttagcggc   20760

taaaaaagtc atggctctgc cctcgggcgg accacgccca tcatgacctt gccaagctcg   20820

tcctgcttct cttcgatctt cgccagcagg gcgaggatcg tggcatcacc gaaccgcgcc   20880

gtgcgcgggt cgtcggtgag ccagagtttc agcaggccgc ccaggcggcc caggtcgcca   20940

ttgatgcggg ccagctcgcg gacgtgctca tagtccacga cgcccgtgat tttgtagccc   21000

tggccgacgg ccagcaggta ggccgacagg ctcatgccgg ccgccgccgc cttttcctca   21060

atcgctcttc gttcgtctgg aaggcagtac accttgatag gtgggctgcc cttcctggtt   21120

ggcttggttt catcagccat ccgcttgccc tcatctgtta cgccggcggt agccggccag   21180

cctcgcagag caggattccc gttgagcacc gccaggtgcg aataagggac agtgaagaag   21240

gaacacccgc tcgcgggtgg gcctacttca cctatcctgc ccggctgacg ccgttggata   21300

caccaaggaa agtctacacg aacccttttgg caaaatcctg tatatcgtgc gaaaaaggat   21360

ggatataccg aaaaaatcgc tataatgacc ccgaagcagg gttatgcagc ggaaaagcgc   21420

tgcttccctg ctgttttgtg gaatatctac cgactggaaa caggcaaatg caggaaatta   21480

ctgaactgag gggacaggcg agagacgatg ccaaagagct acaccgacga gctggccgag   21540

tgggttgaat cccgcgcggc caagaagcgc cggcgtgatg aggctgcggt tgcgttcctg   21600

gcggtgaggg cggatgtcga ggcggcgtta gcgtccggct atgcgctcgt caccatttgg   21660

gagcacatgc gggaaacggg gaaggtcaag ttctcctacg agacgttccg ctcgcacgcc   21720

aggcggcaca tcaaggccaa gcccgccgat gtgcccgcac cgcaggccaa ggctgcggaa   21780

cccgcgccgg cacccaagac gccggagcca cggcggccga agcaggggggg caaggctgaa   21840

aagccggccc ccgctgcggc cccgaccggc ttcaccttca acccaacacc ggacaaaaag   21900

gatctactgt aatggcgaaa attcacatgg ttttgcaggg caaggcgggg tcggcaagt   21960

cggccatcgc cgcgatcatt gcgcagtaca agatggacaa ggggcagaca ccettgtgca   22020

tcgacaccga cccggtgaac gcgacgttcg agggctacaa ggccctgaac gtccgccggc   22080

tgaacatcat ggccggcgac gaaattaact cgcgcaactt cgacaccctg tcgagctga   22140

ttgcgccgac caaggatgac gtggtgatcg acaacggtgc cagctcgttc gtgcctctgt   22200

cgcattacct catcagcaac caggtgccgg ctctgctgca agaaatgggg catgagctgg   22260

tcatccatac cgtcgtcacc ggcggccagg ctctcctgga cacggtgagc ggcttcgccc   22320

agctcgccag ccagttcccg gccgaagcgc ttttcgtggt ctggctgaac ccgtattggg   22380

ggcctatcga gcatgagggc aagagctttg agcagatgaa ggcgtacacg gccaacaagg   22440
```

-continued

```
cccgcgtgtc gtccatcatc cagattccgg ccctcaagga agaaacctac ggccgcgatt    22500

tcagcgacat gctgcaagag cggctgacgt tcgaccaggc gctggccgat gaatcgctca    22560

cgatcatgac gcggcaacgc ctcaagatcg tgcggcgcgg cctgtttgaa cagctcgacg    22620

cggcggccgt gctatgagcg accagattga agagctgatc cgggagattg cggccaagca    22680

cggcatcgcc gtcggccgcg acgacccggt gctgatcctg cataccatca acgcccggct    22740

catggccgac agtgcggcca agcaagagga aatccttgcc gcgttcaagg aagagctgga    22800

agggatcgcc catcgttggg gcgaggacgc caaggccaaa gcggagcgga tgctgaacgc    22860

ggccctggcg gccagcaagg acgcaatggc gaaggtaatg aaggacacgc ccgcgcaggc    22920

ggccgaagcg atccgcaggg aaatcgacga cggccttggc cgccagctcg cggccaaggt    22980

cgcggacgcg cggcgcgtgg cgatgatgaa catgatcgcc ggcggcatgg tgttgttcgc    23040

ggccgccctg gtggtgtggg cctcgttatg aatcgcagag cgcagatga aaaagcccgg    23100

cgttgccggg ctttgttttt gcgttagctg ggcttgtttg acaggcccaa gctctgactg    23160

cgcccgcgct cgcgctcctg ggcctgtttc ttctcctgct cctgcttgcg catcagggcc    23220

tggtgccgtc gggctgcttc acgcatcgaa tcccagtcgc cggccagctc gggatgctcc    23280

gcgcgcatct tgcgcgtcgc cagttcctcg atcttgggcg cgtgaatgcc catgccttcc    23340

ttgatttcgc gcaccatgtc cagccgcgtg tgcaggggtct gcaagcgggc ttgctgttgg    23400

gcctgctgct gctgccaggc ggcctttgta cgcggcaggg acagcaagcc gggggcattg    23460

gactgtagct gctgcaaacg cgcctgctga cggtctacga gctgttctag gcggtcctcg    23520

atgcgctcca cctggtcatg ctttgcctgc acgtagagcg caagggtctg ctggtaggtc    23580

tgctcgatgg gcgcggattc taagagggcc tgctgttccg tctcggcctc ctgggccgcc    23640

tgtagcaaat cctcgccgct gttgccgctg gactgcttta ctgccgggga ctgctgttgc    23700

cctgctcgcg ccgtcgtcgc agttcggctt gcccccactc gattgactgc ttcatttcga    23760

gccgcagcga tgcgatctcg gattgcgtca acggacgggg cagcgcggag gtgtccggct    23820

tctccttggg tgagtcggtc gatgccatag ccaaaggttt ccttccaaaa tgcgtccatt    23880

gctggaccgt gtttctcatt gatgcccgca agcatcttcg gcttgaccgc caggtcaagc    23940

gcgccttcat gggcggtcat gacgacgcc gccatgacct tgccgccgtt gttctcgatg    24000

tagccgcgta atgaggcaat ggtgccgccc atcgtcagcg tgtcatcgac aacgatgtac    24060

ttctggccgg ggatcacctc cccctcgaaa gtcgggttga acgccaggcg atgatctgaa    24120

ccggctccgg ttcgggcgac cttctcccgc tgcacaatgt ccgtttcgac ctcaaggcca    24180

aggcggtcgg ccagaacgac cgccatcatg gccggaatct tgttgttccc cgccgcctcg    24240

acggcgagga ctggaacgat gcggggcttg tcgtcgccga tcagcgtctt gagctgggca    24300

acagtgtcgt ccgaaatcag cgcctcgacc aaattaagcg ccgcttccgc gtcgccctgc    24360

ttcgcagcct ggtattcagg ctcgttggtc aaagaaccaa ggtcgccgtt gcgaaccacc    24420

ttcgggaagt ctccccacgg tgcgcgctcg gctctgctgt agctgctcaa gacgcctccc    24480

tttttagccg ctaaaactct aacgagtgcg cccgcgactc aacttgacgc tttcggcact    24540

tacctgtgcc ttgccacttg cgtcataggt gatgcttttc gcactcccga tttcaggtac    24600

tttatcgaaa tctgaccggg cgtgcattac aaagttcttc cccacctgtt ggtaaatgct    24660

gccgctatct gcgtggacga tgctgccgtc gtggcgctgc gacttatcgg ccttttgggc    24720
```

-continued

```
catatagatg ttgtaaatgc caggtttcag ggccccggct ttatctacct tctggttcgt  24780

ccatgcgcct tggttctcgg tctggacaat tctttgccca ttcatgacca ggaggcggtg  24840

tttcattggg tgactcctga cggttgcctc tggtgttaaa cgtgtcctgg tcgcttgccg  24900

gctaaaaaaa agccgacctc ggcagttcga ggccggcttt ccctagagcc gggcgcgtca  24960

aggttgttcc atctatttta gtgaactgcg ttcgatttat cagttacttt cctcccgctt  25020

tgtgtttcct cccactcgtt tccgcgtcta gccgacccct caacatagcg gcctcttctt  25080

gggctgcctt tgcctcttgc cgcgcttcgt cacgctcggc ttgcaccgtc gtaaagcgct  25140

cggcctgcct ggccgcctct tgcgcgcca acttcctttg ctcctggtgg gcctcggcgt  25200

cggcctgcgc cttcgctttc accgctgcca actccgtgcg caaactctcc gcttcgcgcc  25260

tggtggcgtc gcgctcgccg cgaagcgcct gcatttcctg gttggccgcg tccagggtct  25320

tgcggctctc ttctttgaat gcgcgggcgt cctggtgagc gtagtccagc tcggcgcgca  25380

gctcctgcgc tcgacgctcc acctcgtcgg cccgctgcgt cgccagcgcg gcccgctgct  25440

cggctcctgc cagggcggtg cgtgcttcgg ccagggcttg ccgctggcgt gcggccagct  25500

cggccgcctc ggcggcctgc tgctctagca atgtaacgcg cgcctgggct tcttccagct  25560

cgcgggcctg cgcctcgaag gcgtcggcca gctccccgcg cacggcttcc aactcgttgc  25620

gctcacgatc ccagccggct tgcgctgcct gcaacgattc attggcaagg gcctgggcgg  25680

cttgccagag ggcggccacg gcctggttgc cggcctgctg caccgcgtcc ggcacctgga  25740

ctgccagcgg ggcggcctgc gccgtgcgct ggcgtcgcca ttcgcgcatg ccggcgctgg  25800

cgtcgttcat gttgacgcgg gcggccttac gcactgcatc cacggtcggg aagttctccc  25860

ggtcgccttg ctcgaacagc tcgtccgcag ccgcaaaaat gcggtcgcgc gtctctttgt  25920

tcagttccat gttggctccg gtaattggta agaataataa tactcttacc taccttatca  25980

gcgcaagagt ttagctgaac agttctcgac ttaacggcag gttttttagc ggctgaaggg  26040

caggcaaaaa aagccccgca cggtcggcgg gggcaaaggg tcagcgggaa ggggattagc  26100

gggcgtcggg cttcttcatg cgtcggggcc gcgcttcttg ggatggagca cgacgaagcg  26160

cgcacgcgca tcgtcctcgg ccctatcggc ccgcgtcgcg gtcaggaact tgtcgcgcgc  26220

taggtcctcc ctggtgggca ccaggggcat gaactcggcc tgctcgatgt aggtccactc  26280

catgaccgca tcgcagtcga ggccgcgttc cttcaccgtc tcttgcaggt cgcggtacgc  26340

ccgctcgttg agcggctggt aacgggccaa ttggtcgtaa atggctgtcg gccatgagcg  26400

gcctttcctg ttgagccagc agccgacgac gaagccggca atgcaggccc ctggcacaac  26460

caggccgacg ccgggggcag gggatggcag cagctcgcca accaggaacc ccgccgcgat  26520

gatgccgatg ccggtcaacc agcccttgaa actatccggc cccgaaacac ccctgcgcat  26580

tgcctggatg ctgcgccgga tagcttgcaa catcaggagc cgtttctttt gttcgtcagt  26640

catggtccgc cctcaccagt tgttcgtatc ggtgtcggac gaactgaaat cgcaagagct  26700

gccggtatcg gtccagccgc tgtccgtgtc gctgctgccg aagcacggcg aggggtccgc  26760

gaacgccgca gacggcgtat ccggccgcag cgcatcgccc agcatggccc cggtcagcga  26820

gccgccggcc aggtagccca gcatggtgct gttggtcgcc ccggccacca gggccgacgt  26880

gacgaaatcg ccgtcattcc ctctggattg ttcgctgctc ggcggggcag tgcgccgcgc  26940

cggcggcgtc gtggatggct cgggttggct ggcctgcgac ggccggcgaa aggtgcgcag  27000
```

-continued

```
cagctcgtta tcgaccggct gcggcgtcgg ggccgccgcc ttgcgctgcg gtcggtgttc    27060

cttcttcggc tcgcgcagct tgaacagcat gatcgcggaa accagcagca acgccgcgcc    27120

tacgcctccc gcgatgtaga acagcatcgg attcattctt cggtcctcct tgtagcggaa    27180

ccgttgtctg tgcggcgcgg gtggcccgcg ccgctgtctt tggggatcag ccctcgatga    27240

gcgcgaccag tttcacgtcg gcaaggttcg cctcgaactc ctggccgtcg tcctcgtact    27300

tcaaccaggc atagccttcc gccggcggcc gacggttgag gataaggcgg gcagggcgct    27360

cgtcgtgctc gacctggacg atggcctttt tcagcttgtc cgggtccggc tccttcgcgc    27420

cctttccctt ggcgtcctta ccgtcctggt cgccgtcctc gccgtcctgg ccgtcgccgg    27480

cctccgcgtc acgctcggca tcagtctggc cgttgaaggc atcgacggtg ttgggatcgc    27540

ggccttctc gtccaggaac tcgcgcagca gcttgaccgt gccgcgcgtg atttcctggg     27600

tgtcgtcgtc aagccacgcc tcgacttcct ccgggcgctt cttgaaggcc gtcaccagct    27660

cgttcaccac ggtcacgtcg cgcacgcggc cggtgttgaa cgcatcggcg atcttctccg    27720

gcaggtccag cagcgtgacg tgctgggtga tgaacgccgg cgacttgccg atttccttgg    27780

cgatatcgcc tttcttcttg cccttcgcca gctcgcggcc aatgaagtcg gcaatttcgc    27840

gcggggtcag ctcgttgcgt tgcaggttct cgataacctg gtcggcttcg ttgtagtcgt    27900

tgtcgatgaa cgccgggatg gacttcttgc cggcccactt cgagccacgg tagcggcggg    27960

cgccgtgatt gatgatatag cggcccggct gctcctggtt ctcgcgcacc gaaatgggtg    28020

acttcacccc gcgctctttg atcgtggcac cgatttccgc gatgctctcc ggggaaaagc    28080

cggggttgtc ggccgtccgc ggctgatgcg gatcttcgtc gatcaggtcc aggtccagct    28140

cgatagggcc ggaaccgccc tgagacgccg caggagcgtc caggaggctc gacaggtcgc    28200

cgatgctatc caaccccagg ccggacggct gcgccgcgcc tgcggcttcc tgagcggccg    28260

cagcggtgtt tttcttggtg gtcttggctt gagccgcagt cattgggaaa tctccatctt    28320

cgtgaacacg taatcagcca gggcgcgaac ctctttcgat gccttgcgcg cggccgtttt    28380

cttgatcttc cagaccggca caccggatgc gagggcatcg gcgatgctgc tgcgcaggcc    28440

aacggtggcc ggaatcatca tcttggggta cgcggccagc agctcggctt ggtggcgcgc    28500

gtggcgcgga ttccgcgcat cgaccttgct gggcaccatg ccaaggaatt gcagcttggc    28560

gttcttctgg cgcacgttcg caatggtcgt gaccatcttc ttgatgccct ggatgctgta    28620

cgcctcaagc tcgatggggg acagcacata gtcggccgcg aagagggcgg ccgccaggcc    28680

gacgccaagg gtcggggccg tgtcgatcag gcacacgtcg aagccttggt cgcgcagggc    28740

cttgatgttc gccccgaaca gctcgcgggc gtcgtccagc gacagccgtt cggcgttcgc    28800

cagtaccggg ttggactcga tgagggcgag gcgcgcggcc tggccgtcgc cggctgcggg    28860

tgcggtttcg gtccagccgc cggcagggac agcgccgaac agcttgcttg catgcaggcc    28920

ggtagcaaag tccttgagcg tgtaggacgc attgccctgg gggtccaggt cgatcacggc    28980

aacccgcaag ccgcgctcga aaaagtcgaa ggcaagatgc acaagggtcg aagtcttgcc    29040

gacgccgcct ttctggttgg ccgtgaccaa agttttcatc gtttggtttc ctgttttttc    29100

ttggcgtccg cttcccactt ccggacgatg tacgcctgat gttccggcag aaccgccgtt    29160

acccgcgcgt accctcgggg caagttcttg tcctcgaacg cggcccacac gcgatgcacc    29220

gcttgcgaca ctgcgcccct ggtcagtccc agcgacgttg cgaacgtcgc ctgtggcttc    29280
```

-continued

```
ccatcgacta agacgccccg cgctatctcg atggtctgct gccccacttc cagcccctgg    29340

atcgcctcct ggaactggct ttcggtaagc cgtttcttca tggataacac ccataatttg    29400

ctccgcgcct tggttgaaca tagcggtgac agccgccagc acatgagaga agtttagcta    29460

aacatttctc gcacgtcaac acctttagcc gctaaaactc gtccttggcg taacaaaaca    29520

aaagcccgga aaccgggctt tcgtctcttg ccgcttatgg ctctgcaccc ggctccatca    29580

ccaacaggtc gcgcacgcgc ttcactcggt tgcggatcga cactgccagc ccaacaaagc    29640

cggttgccgc cgccgccagg atcgcgccga tgatgccggc cacaccggcc atcgcccacc    29700

aggtcgcgc cttccggttc cattcctgct ggtactgctt cgcaatgctg acctcggct      29760

caccataggc tgaccgctcg atggcgtatg ccgcttctcc ccttggcgta aaacccagcg    29820

ccgcaggcgg cattgccatg ctgcccgccg ctttcccgac cacgacgcgc gcaccaggct    29880

tgcggtccag accttcggcc acggcgagct gcgcaaggac ataatcagcc gccgacttgg    29940

ctccacgcgc ctcgatcagc tcttgcactc gcgcgaaatc cttggcctcc acggccgcca    30000

tgaatcgcgc acgcggcgaa ggctccgcag ggccggcgtc gtgatcgccg ccgagaatgc    30060

ccttcaccaa gttcgacgac acgaaaatca tgctgacggc tatcaccatc atgcagacgg    30120

atcgcacgaa cccgctgaat tgaacacgag cacggcaccc gcgaccacta tgccaagaat    30180

gcccaaggta aaaattgccg gccccgccat gaagtccgtg aatgccccga cggccgaagt    30240

gaagggcagg ccgccaccca ggccgccgcc ctcactgccc ggcacctggt cgctgaatgt    30300

cgatgccagc acctgcggca cgtcaatgct tccgggcgtc gcgctcgggc tgatcgccca    30360

tcccgttact gccccgatcc cggcaatggc aaggactgcc agcgctgcca tttttggggt    30420

gaggccgttc gcggccgagg ggcgcagccc ctggggggat gggaggcccg cgttagcggg    30480

ccgggagggt tcgagaaggg ggggcacccc ccttcggcgt gcgcggtcac gcgcacaggg    30540

cgcagccctg gttaaaaaca aggtttataa atattggttt aaaagcaggt taaaagacag    30600

gttagcggtg gccgaaaaac gggcggaaac ccttgcaaat gctggatttt ctgcctgtgg    30660

acagcccctc aaatgtcaat aggtgcgccc ctcatctgtc agcactctgc ccctcaagtg    30720

tcaaggatcg cgcccctcat ctgtcagtag tcgcgcccct caagtgtcaa taccgcaggg    30780

cacttatccc caggcttgtc cacatcatct gtgggaaact cgcgtaaaat caggcgtttt    30840

cgccgatttg cgaggctggc cagctccacg tcgccggccg aaatcgagcc tgcccctcat    30900

ctgtcaacgc cgcgccgggt gagtcggccc ctcaagtgtc aacgtccgcc cctcatctgt    30960

cagtgagggc caagttttcc gcgaggtatc cacaacgccg cgcggccgcgg tgtctcgcac    31020

acggcttcga cggcgtttct ggcgcgtttg cagggccata gacggccgcc agcccagcgg    31080

cgagggcaac cagcccggtg agcgtcggaa aggcgctgga agcccgtag cgacgcggag     31140

aggggcgaga caagccaagg gcgcaggctc gatgcgcagc acgacatagc cggttctcgc    31200

aaggacgaga atttccctgc ggtgcccctc aagtgtcaat gaaagtttcc aacgcgagcc    31260

attcgcgaga gccttgagtc cacgctagat gagagctttg ttgtaggtgg accagttggt    31320

gattttgaac ttttgctttg ccacggaacg gtctgcgttg tcgggaagat gcgtgatctg    31380

atccttcaac tcagcaaaag ttcgatttat tcaacaaagc cacgttgtgt ctcaaaatct    31440

ctgatgttac attgcacaag ataaaaatat atcatcatga acaataaaac tgtctgctta    31500

cataaacagt aatacaaggg gtgttatgag ccatattcaa cgggaaacgt cttgctcgac    31560
```

-continued

```
tctagagctc gttcctcgag gaacggtacc tgcggggaag cttacaataa tgtgtgttgt    31620

taagtcttgt tgcctgtcat cgtctgactg actttcgtca taaatcccgg cctccgtaac    31680

ccagctttgg gcaagctcac ggatttgatc cggcggaacg ggaatatcga gatgccgggc    31740

tgaacgctgc agttccagct ttccctttcg ggacaggtac tccagctgat tgattatctg    31800

ctgaagggtc ttggttccac ctcctggcac aatgcgaatg attacttgag cgcgatcggg    31860

catccaattt tctcccgtca ggtgcgtggt caagtgctac aaggcacctt tcagtaacga    31920

gcgaccgtcg atccgtcgcc gggatacgga caaaatggag cgcagtagtc catcgagggc    31980

ggcgaaagcc tcgccaaaag caatacgttc atctcgcaca gcctccagat ccgatcgagg    32040

gtcttcggcg taggcagata gaagcatgga tacattgctt gagagtattc cgatggactg    32100

aagtatggct tccatctttt ctcgtgtgtc tgcatctatt tcgagaaagc ccccgatgcg    32160

gcgcaccgca acgcgaattg ccatactatc cgaaagtccc agcaggcgcg cttgatagga    32220

aaaggtttca tactcggccg atcgcagacg ggcactcacg accttgaacc cttcaacttt    32280

cagggatcga tgctggttga tggtagtctc actcgacgtg gctctggtgt gttttgacat    32340

agcttcctcc aaagaaagcg gaaggtctgg atactccagc acgaaatgtg cccgggtaga    32400

cggatggaag tctagccctg ctcaatatga aatcaacagt acatttacag tcaatactga    32460

atatacttgc tacatttgca attgtcttat aacgaatgtg aaataaaaat agtgtaacaa    32520

cgcttttact catcgataat cacaaaaaca tttatacgaa caaaaataca aatgcactcc    32580

ggtttcacag gataggcggg atcagaatat gcaacttttg acgttttgtt ctttcaaagg    32640

gggtgctggc aaaaccaccg cactcatggg cctttgcgct gctttggcaa atgacggtaa    32700

acgagtggcc ctctttgatg ccgacgaaaa ccggcctctg acgcgatgga gagaaaacgc    32760

cttacaaagc agtactggga tcctcgctgt gaagtctatt ccgccgacga aatgcccctt    32820

cttgaagcag cctatgaaaa tgccgagctc gaaggatttg attatgcgtt ggccgatacg    32880

cgtggcggct cgagcgagct caacaacaca atcatcgcta gctcaaacct gcttctgatc    32940

cccaccatgc taacgccgct cgacatcgat gaggcactat ctacctaccg ctacgtcatc    33000

gagctgctgt tgagtgaaaa tttggcaatt cctacagctg ttttgcgcca acgcgtcccg    33060

gtcggccgat tgacaacatc gcaacgcagg atgtcagaga cgctagagag ccttccagtt    33120

gtaccgtctc ccatgcatga aagagatgca tttgccgcga tgaaagaacg cggcatgttg    33180

catcttacat tactaaacac gggaactgat ccgacgatgc gcctcataga gaggaatctt    33240

cggattgcga tggaggaagt cgtggtcatt tcgaaactga tcagcaaaat cttggaggct    33300

tgaagatggc aattcgcaag cccgcattgt cggtcggcga agcacggcgg cttgctggtg    33360

ctcgacccga gatccaccat cccaacccga cacttgttcc ccagaagctg gacctccagc    33420

acttgcctga aaaagccgac gagaaagacc agcaacgtga gcctctcgtc gccgatcaca    33480

tttacagtcc cgatcgacaa cttaagctaa ctgtggatgc ccttagtcca cctccgtccc    33540

cgaaaaagct ccaggttttt ctttcagcgc gaccgcccgc gcctcaagtg tcgaaaacat    33600

atgacaacct cgttcggcaa tacagtccct cgaagtcgct acaaatgatt ttaaggcgcg    33660

cgttggacga tttcgaaagc atgctggcag atggatcatt tcgcgtggcc ccgaaaagtt    33720

atccgatccc ttcaactaca gaaaaatccg ttctcgttca gacctcacgc atgttcccgg    33780

ttgcgttgct cgaggtcgct cgaagtcatt ttgatccgtt ggggttggag accgctcgag    33840
```

-continued

```
ctttcggcca caagctggct accgccgcgc tcgcgtcatt ctttgctgga gagaagccat    33900

cgagcaattg gtgaagaggg acctatcgga accccctcacc aaatattgag tgtaggtttg    33960

aggccgctgg ccgcgtcctc agtcacctttt tgagccagat aattaagagc caaatgcaat    34020

tggctcaggc tgccatcgtc cccccgtgcg aaacctgcac gtccgcgtca aagaaataac    34080

cggcacctct tgctgttttt atcagttgag ggcttgacgg atccgcctca agtttgcggc    34140

gcagccgcaa aatgagaaca tctatactcc tgtcgtaaac ctcctcgtcg cgtactcgac    34200

tggcaatgag aagttgctcg cgcgatagaa cgtcgcgggg tttctctaaa aacgcgagga    34260

gaagattgaa ctcacctgcc gtaagtttca cctcaccgcc agcttcggac atcaagcgac    34320

gttgcctgag attaagtgtc cagtcagtaa aacaaaaaga ccgtcggtct ttggagcgga    34380

caacgttggg gcgcacgcgc aaggcaaccc gaatgcgtgc aagaaactct ctcgtactaa    34440

acggcttagc gataaaatca cttgctccta gctcgagtgc aacaacttta tccgtctcct    34500

caaggcggtc gccactgata attatgattg gaatatcaga ctttgccgcc agatttcgaa    34560

cgatctcaag cccatcttca cgacctaaat ttagatcaac aaccacgaca tcgaccgtcg    34620

cggaagagag tactctagtg aactgggtgc tgtcggctac cgcggtcact ttgaaggcgt    34680

ggatcgtaag gtattcgata ataagatgcc gcatagcgac atcgtcatcg ataagaagaa    34740

cgtgtttcaa cggctcacct ttcaatctaa aatctgaacc cttgttcaca gcgcttgaga    34800

aattttcacg tgaaggatgt acaatcatct ccagctaaat gggcagttcg tcagaattgc    34860

ggctgaccgc ggatgacgaa aatgcgaacc aagtatttca attttatgac aaaagttctc    34920

aatcgttgtt acaagtgaaa cgcttcgagg ttacagctac tattgattaa ggagatcgcc    34980

tatggtctcg ccccggcgtc gtgcgtccgc cgcgagccag atctcgccta cttcataaac    35040

gtcctcatag gcacggaatg gaatgatgac atcgatcgcc gtagagagca tgtcaatcag    35100

tgtgcgatct tccaagctag caccttgggc gctacttttg acaagggaaa acagtttctt    35160

gaatccttgg attggattcg cgccgtgtat tgttgaaatc gatcccggat gtcccgagac    35220

gacttcactc agataagccc atgctgcatc gtcgcgcatc tcgccaagca atatccggtc    35280

cggccgcata cgcagacttg cttggagcaa gtgctcggcg ctcacagcac ccagcccagc    35340

accgttcttg gagtagagta gtctaacatg attatcgtgt ggaatgacga gttcgagcgt    35400

atcttctatg gtgattagcc tttcctgggg ggggatggcg ctgatcaagg tcttgctcat    35460

tgttgtcttg ccgcttccgg tagggccaca tagcaacatc gtcagtcggc tgacgacgca    35520

tgcgtgcaga aacgcttcca aatccccgtt gtcaaaatgc tgaaggatag cttcatcatc    35580

ctgattttgg cgtttccttc gtgtctgcca ctggttccac ctcgaagcat cataacggga    35640

ggagacttct ttaagaccag aaaacacgcga gcttggccgt cgaatggtca agctgacggt    35700

gcccgaggga acggtcggcg gcagacagat ttgtagtcgt tcaccaccag gaagttcagt    35760

ggcgcagagg gggttacgtg gtccgacatc ctgctttctc agcgcgcccg ctaaaatagc    35820

gatatcttca agatcatcat aagagacggg caaaggcatc ttggtaaaaa tgccggcttg    35880

gcgcacaaat gcctctccag gtcgattgat cgcaatttct tcagtcttcg ggtcatcgag    35940

ccattccaaa atcggcttca gaagaaagcg tagttgcgga tccacttcca tttacaatgt    36000

atcctatctc taagcggaaa tttgaattca ttaagagcgg cggttcctcc cccgcgtggc    36060

gccgccagtc aggcggagct ggtaaacacc aaagaaatcg aggtcccgtg ctacgaaaat    36120
```

-continued

```
ggaaacggtg tcaccctgat tcttcttcag ggttggcggt atgttgatgg ttgccttaag    36180

ggctgtctca gttgtctgct caccgttatt ttgaaagctg ttgaagctca tcccgccacc    36240

cgagctgccg gcgtaggtgc tagctgcctg gaaggcgcct tgaacaacac tcaagagcat    36300

agctccgcta aaacgctgcc agaagtggct gtcgaccgag cccggcaatc ctgagcgacc    36360

gagttcgtcc gcgcttggcg atgttaacga gatcatcgca tggtcaggtg tctcggcgcg    36420

atcccacaac acaaaaacgc gcccatctcc ctgttgcaag ccacgctgta tttcgccaac    36480

aacggtggtg ccacgatcaa gaagcacgat attgttcgtt gttccacgaa tatcctgagg    36540

caagacacac tttacatagc ctgccaaatt tgtgtcgatt gcggtttgca agatgcacgg    36600

aattattgtc ccttgcgtta ccataaaatc ggggtgcggc aagagcgtgg cgctgctggg    36660

ctgcagctcg gtgggtttca tacgtatcga caaatcgttc tcgccggaca cttcgccatt    36720

cggcaaggag ttgtcgtcac gcttgccttc ttgtcttcgg cccgtgtcgc cctgaatggc    36780

gcgtttgctg acccttgat cgccgctgct atatgcaaaa atcggtgttt cttccggccg    36840

tggctcatgc cgctccggtt cgcccctcgg cggtagagga gcagcaggct gaacagcctc    36900

ttgaaccgct ggaggatccg gcggcacctc aatcggagct ggatgaaatg gcttggtgtt    36960

tgttgcgatc aaagttgacg gcgatgcgtt ctcattcacc ttcttttggc gcccacctag    37020

ccaaatgagg cttaatgata acgcgagaac gacacctccg acgatcaatt tctgagaccc    37080

cgaaagacgc cggcgatgtt gtcggagac cagggatcca gatgcatcaa cctcatgtgc    37140

cgcttgctga ctatcgttat tcatcccttc gccccttca ggacgcgttt cacatcgggc    37200

ctcaccgtgc ccgtttgcgg cctttggcca acgggatcgt aagcggtgtt ccagatacat    37260

agtactgtgt ggccatccct cagacgccaa cctcgggaaa ccgaagaaat ctcgacatcg    37320

ctcccttta ctgaatagtt ggcaacagct tccttgccat caggattgat ggtgtagatg    37380

gagggtatgc gtacattgcc cggaaagtgg aataccgtcg taaatccatt gtcgaagact    37440

tcgagtggca acagcgaacg atcgccttgg gcgacgtagt gccaattact gtccgccgca    37500

ccaagggctg tgacaggctg atccaataaa ttctcagctt tccgttgata ttgtgcttcc    37560

gcgtgtagtc tgtccacaac agccttctgt tgtgcctccc ttcgccgagc cgccgcatcg    37620

tcggcgggt aggcgaattg gacgctgtaa tagagatcgg ctgctcttt atcgaggtgg    37680

gacagagtct tggaacttat actgaaaaca taacggcgca tcccgagtc gcttgcggtt    37740

agcacgatta ctggctgagg cgtgaggacc tggcttgcct tgaaaaatag ataatttccc    37800

cgcggtaggg ctgctagatc tttgctattt gaaacggcaa ccgctgtcac cgtttcgttc    37860

gtggcgaatg ttacgaccaa agtagctcca accgccgtcg agaggcgcac cacttgatcg    37920

ggattgtaag ccaaataacg catgcgcgga tctagcttgc ccgccattgg agtgtcttca    37980

gcctccgcac cagtcgcagc ggcaaataaa catgctaaaa tgaaaagtgc ttttctgatc    38040

atggttcgct gtggcctacg tttgaaacgg tatcttccga tgtctgatag gaggtgacaa    38100

ccagacctgc cgggttggtt agtctcaatc tgccgggcaa gctggtcacc ttttcgtagc    38160

gaactgtcgc ggtccacgta ctcaccacag gcattttgcc gtcaacgacg agggtccttt    38220

tatagcgaat ttgctgcgtg cttggagtta catcatttga agcgatgtgc tcgacctcca    38280

ccctgccgcg tttgccaaga atgacttgag gcgaactggg attgggatag ttgaagaatt    38340

gctggtaatc ctggcgcact gttggggcac tgaagttcga taccaggtcg taggcgtact    38400
```

-continued

```
gagcggtgtc ggcatcataa ctctcgcgca ggcgaacgta ctcccacaat gaggcgttaa    38460

cgacggcctc ctcttgagtt gcaggcaatc gcgagacaga cacctcgctg tcaacggtgc    38520

cgtccggccg tatccataga tatacgggca caagcctgct caacggcacc attgtggcta    38580

tagcgaacgc ttgagcaaca tttcccaaaa tcgcgatagc tgcgacagct gcaatgagtt    38640

tggagagacg tcgcgccgat ttcgctcgcg cggtttgaaa ggcttctact tccttatagt    38700

gctcggcaag gctttcgcgc gccactagca tggcatattc aggccccgtc atagcgtcca    38760

cccgaattgc cgagctgaag atctgacgga gtaggctgcc atcgcccac attcagcggg     38820

aagatcgggc ctttgcagct cgctaatgtg tcgtttgtct ggcagccgct caaagcgaca    38880

actaggcaca gcaggcaata cttcatagaa ttctccattg aggcgaattt ttgcgcgacc    38940

tagcctcgct caacctgagc gaagcgacgg tacaagctgc tggcagattg ggttgcgccg    39000

ctccagtaac tgcctccaat gttgccggcg atcgccggca aagcgacaat gagcgcatcc    39060

cctgtcagaa aaaacatatc gagttcgtaa agaccaatga tcttggccgc ggtcgtaccg    39120

gcgaaggtga ttacaccaag cataagggtg agcgcagtcg cttcggttag gatgacgatc    39180

gttgccacga ggtttaagag gagaagcaag agaccgtagg tgataagttg cccgatccac    39240

ttagctgcga tgtcccgcgt gcgatcaaaa atatatccga cgaggatcag aggcccgatc    39300

gcgagaagca ctttcgtgag aattccaacg gcgtcgtaaa ctccgaaggc agaccagagc    39360

gtgccgtaaa ggacccactg tgccccttgg aaagcaagga tgtcctggtc gttcatcgga    39420

ccgattcgg atgcgatttt ctgaaaaacg gcctgggtca cggcgaacat tgtatccaac     39480

tgtgccggaa cagtctgcag aggcaagccg gttacactaa actgctgaac aaagtttggg    39540

accgtctttt cgaagatgga aaccacatag tcttggtagt tagcctgccc aacaattaga    39600

gcaacaacga tggtgaccgt gatcacccga gtgataccgc tacgggtatc gacttcgccg    39660

cgtatgacta aaataccctg aacaataatc caaagagtga cacaggcgat caatggcgca    39720

ctcaccgcct cctggatagt ctcaagcatc gagtccaagc ctgtcgtgaa ggctacatcg    39780

aagatcgtat gaatggccgt aaacggcgcc ggaatcgtga aattcatcga ttggacctga    39840

acttgactgg tttgtcgcat aatgttggat aaaatgagct cgcattcggc gaggatgcgg    39900

gcggatgaac aaaatcgccca gcccttagggg agggcaccaa agatgacagc ggtcttttga    39960

tgctccttgc gttgagcggc cgcctcttcc gcctcgtgaa ggccggcctg cgcggtagtc    40020

atcgttaata ggcttgtcgc ctgtacattt tgaatcattg cgtcatggat ctgcttgaga    40080

agcaaaccat tggtcacggt tgcctgcatg atattgcgag atcgggaaag ctgagcagac    40140

gtatcagcat tcgccgtcaa gcgtttgtcc atcgtttcca gattgtcagc cgcaatgcca    40200

gcgctgtttg cggaaccggt gatctgcgat cgcaacaggt ccgcttcagc atcactaccc    40260

acgactgcac gatctgtatc gctggtgatc gcacgtgccg tggtcgacat tggcattcgc    40320

ggcgaaaaca tttcattgtc taggtccttc gtcgaaggat actgattttt ctggttgagc    40380

gaagtcagta gtccagtaac gccgtaggcc gacgtcaaca tcgtaaccat cgctatagtc    40440

tgagtgagat tctccgcagt cgcgagcgca gtcgcgagcg tctcagcctc cgttgccggg    40500

tcgctaacaa caaactgcgc ccgcgcgggc tgaatatata gaaagctgca ggtcaaaact    40560

gttgcaataa gttgcgtcgt cttcatcgtt tcctacctta tcaatcttct gcctcgtggt    40620

gacgggccat gaattcgctg agccagccag atgagttgcc ttcttgtgcc tcgcgtagtc    40680
```

-continued

```
gagttgcaaa gcgcaccgtg ttggcacgcc ccgaaagcac ggcgacatat tcacgcatat    40740

cccgcagatc aaattcgcag atgacgcttc cactttctcg tttaagaaga aacttacggc    40800

tgccgaccgt catgtcttca cggatcgcct gaaattcctt ttcggtacat ttcagtccat    40860

cgacataagc cgatcgatct gcggttggtg atggatagaa aatcttcgtc atacattgcg    40920

caaccaagct ggctcctagc ggcgattcca gaacatgctc tggttgctgc gttgccagta    40980

ttagcatccc gttgtttttt cgaacggtca ggaggaattt gtcgacgaca gtcgaaaatt    41040

tagggtttaa caaataggcg cgaaactcat cgcagctcat cacaaaacgg cggccgtcga    41100

tcatggctcc aatccgatgc aggagatatg ctgcagcggg agcgcatact tcctcgtatt    41160

cgagaagatg cgtcatgtcg aagccggtaa tcgacggatc taactttact tcgtcaactt    41220

cgccgtcaaa tgcccagcca agcgcatggc cccggcacca gcgttggagc cgcgctcctg    41280

cgccttcggc gggcccatgc aacaaaaatt cacgtaaccc cgcgattgaa cgcatttgtg    41340

gatcaaacga gagctgacga tggataccac ggaccagacg gcggttctct tccggagaaa    41400

tcccaccccg accatcactc tcgatgagag ccacgatcca ttcgcgcaga aaatcgtgtg    41460

aggctgctgt gttttctagg ccacgcaacg gcgccaaccc gctgggtgtg cctctgtgaa    41520

gtgccaaata tgttcctcct gtggcgcgaa ccagcaattc gccaccccgg tccttgtcaa    41580

agaacacgac cgtacctgca cggtcgacca tgctctgttc gagcatggct agaacaaaca    41640

tcatgagcgt cgtcttaccc ctcccgatag gcccgaatat tgccgtcatg ccaacatcgt    41700

gctcatgcgg gatatagtcg aaaggcgttc cgccattggt acgaaatcgg gcaatcgcgt    41760

tgccccagtg gcctgagctg gcgccctctg gaaagttttc gaaagagaca aaccctgcga    41820

aattgcgtga agtgattgcg ccagggcgtg tgcgccactt aaaattcccc ggcaattggg    41880

accaataggc cgcttccata ccaatacctt cttggacaac cacggcacct gcatccgcca    41940

ttcgtgtccg agcccgcgcg cccctgtccc caagactatt gagatcgtct gcatagacgc    42000

aaaggctcaa atgatgtgag cccataacga attcgttgct cgcaagtgcg tcctcagcct    42060

cggataattt gccgatttga gtcacggctt tatcgccgga actcagcatc tggctcgatt    42120

tgaggctaag tttcgcgtgc gcttgcgggc gagtcaggaa cgaaaaactc tgcgtgagaa    42180

caagtggaaa atcgagggat agcagcgcgt tgagcatgcc cggccgtgtt tttgcagggt    42240

attcgcgaaa cgaatagatg gatccaacgt aactgtcttt tggcgttctg atctcgagtc    42300

ctcgcttgcc gcaaatgact ctgtcggtat aaatcgaagc gccgagtgag ccgctgacga    42360

ccggaaccgg tgtgaaccga ccagtcatga tcaaccgtag cgcttcgcca atttcggtga    42420

agagcacacc ctgcttctcg cggatgccaa gacgatgcag gccatacgct ttaagagagc    42480

cagcgacaac atgccaaaga tcttccatgt tcctgatctg gcccgtgaga tcgtttccc     42540

tttttccgct tagcttggtg aacctcctct ttaccttccc taaagccgcc tgtgggtaga    42600

caatcaacgt aaggaagtgt tcattgcgga ggagttggcc ggagagcacg cgctgttcaa    42660

aagcttcgtt caggctagcg gcgaaaacac tacggaagtg tcgcggcgcc gatgatggca    42720

cgtcggcatg acgtacgagg tgagcatata ttgacacatg atcatcagcg atattgcgca    42780

acagcgtgtt gaacgcacga caacgcgcat tgcgcatttc agtttcctca agctcgaatg    42840

caacgccatc aattctcgca atggtcatga tcgatccgtc ttcaagaagg acgatatggt    42900

cgctgaggtg gccaatataa gggagataga tctcaccgga tctttcggtc gttccactcg    42960
```

-continued

```
cgccgagcat cacaccattc ctctccctcg tgggggaacc ctaattggat ttgggctaac   43020

agtagcgccc ccccaaactg cactatcaat gcttcttccc gcggtccgca aaaatagcag   43080

gacgacgctc gccgcattgt agtctcgctc cacgatgagc cgggctgcaa accataacgg   43140

cacgagaacg acttcgtaga gcgggttctg aacgataacg atgacaaagc cggcgaacat   43200

catgaataac cctgccaatg tcagtggcac cccaagaaac aatgcgggcc gtgtggctgc   43260

gaggtaaagg gtcgattctt ccaaacgatc agccatcaac taccgccagt gagcgtttgg   43320

ccgaggaagc tcgccccaaa catgataaca atgccgccga cgacgccggc aaccagccca   43380

agcgaagccc gcccgaacat ccaggagatc ccgatagcga caatgccgag aacagcgagt   43440

gactggccga acggaccaag gataaacgtg catatattgt taaccattgt ggcggggtca   43500

gtgccgccac ccgcagattg cgctgcggcg ggtccggatg aggaaatgct ccatgcaatt   43560

gcaccgcaca agcttggggc gcagctcgat atcacgcgca tcatcgcatt cgagagcgag   43620

aggcgattta gatgtaaacg gtatctctca aagcatcgca tcaatgcgca cctccttagt   43680

ataagtcgaa taagacttga ttgtcgtctg cggatttgcc gttgtcctgg tgtggcggtg   43740

gcggagcgat taaaccgcca gcgccatcct cctgcgagcg gcgctgatat gaccccccaaa  43800

catcccacgt ctcttcggat tttagcgcct cgtgatcgtc ttttggaggc tcgattaacg   43860

cgggcaccag cgattgagca gctgtttcaa cttttcgcac gtagccgttt gcaaaaccgc   43920

cgatgaaatt accggtgttg taagcggaga tcgcccgacg aagcgcaaat tgcttctcgt   43980

caatcgtttc gccgcctgca taacgacttt tcagcatgtt tgcagcggca gataatgatg   44040

tgcacgcctg gagcgcaccg tcaggtgtca gaccgagcat agaaaaattt cgagagttta   44100

tttgcatgag gccaacatcc agcgaatgcc gtgcatcgag acggtgcctg acgacttggg   44160

ttgcttggct gtgatcttgc cagtgaagcg tttcgccggt cgtgttgtca tgaatcgcta   44220

aaggatcaaa gcgactctcc accttagcta tcgccgcaag cgtagatgtc gcaactgatg   44280

gggcacactt gcgagcaaca tggtcaaact cagcagatga gagtggcgtg gcaaggctcg   44340

acgaacagaa ggagaccatc aaggcaagag aaagcgaccc cgatctctta agcatacctt   44400

atctccttag ctcgcaacta caccgcctc tcccgttgga agaagtgcgt tgttttatgt    44460

tgaagattat cgggagggtc ggttactcga aaattttcaa ttgcttcttt atgatttcaa   44520

ttgaagcgag aaacctcgcc cggcgtcttg gaacgcaaca tggaccgaga accgcgcatc   44580

catgactaag caaccggatc gacctattca ggccgcagtt ggtcaggtca ggctcagaac   44640

gaaaatgctc ggcgaggtta cgctgtctgt aaacccattc gatgaacggg aagcttcctt   44700

ccgattgctc ttggcaggaa tattggccca tgcctgcttg cgctttgcaa atgctcttat   44760

cgcgttggta tcatatgcct tgtccgccag cagaaacgca ctctaagcga ttatttgtaa   44820

aaatgtttcg gtcatgcggc ggtcatgggc ttgacccgct gtcagcgcaa gacggatcgg   44880

tcaaccgtcg gcatcgacaa cagcgtgaat cttggtggtc aaaccgccac gggaacgtcc   44940

catacagcca tcgtcttgat cccgctgttt cccgtcgccg catgttggtg gacgcggaca   45000

caggaactgt caatcatgac gacattctat cgaaagcctt ggaaatcaca ctcagaatat   45060

gatcccagac gtctgcctca cgccatcgta caaagcgatt gtagcaggtt gtacaggaac   45120

cgtatcgatc aggaacgtct gcccagggcg ggccgtccg gaagcgccac aagatgacat    45180

tgatcacccg cgtcaacgcg cggcacgcga cgcggcttat ttgggaacaa aggactgaac   45240
```

```
aacagtccat tcgaaatcgg tgacatcaaa gcggggacgg gttatcagtg gcctccaagt    45300

caagcctcaa tgaatcaaaa tcagaccgat ttgcaaacct gatttatgag tgtgcggcct    45360

aaatgatgaa atcgtccttc tagatcgcct ccgtggtgta gcaacacctc gcagtatcgc    45420

cgtgctgacc ttggccaggg aattgactgg caagggtgct ttcacatgac cgctcttttg    45480

gccgcgatag atgatttcgt tgctgctttg ggcacgtaga aggagagaag tcatatcgga    45540

gaaattcctc ctggcgcgag agcctgctct atcgcgacgg catcccactg tcgggaacag    45600

accggatcat tcacgaggcg aaagtcgtca acacatgcgt tataggcatc ttcccttgaa    45660

ggatgatctt gttgctgcca atctggaggt gcggcagccg caggcagatg cgatctcagc    45720

gcaacttgcg gcaaaacatc tcactcacct gaaaaccact agcgagtctc gcgatcagac    45780

gaaggccttt tacttaacga cacaatatcc gatgtctgca tcacaggcgt cgctatccca    45840

gtcaatacta aagcggtgca ggaactaaag attactgatg acttaggcgt gccacgaggc    45900

ctgagacgac gcgcgtagac agttttttga aatcattatc aaagtgatgg cctccgctga    45960

agcctatcac ctctgcgccg gtctgtcgga gagatgggca agcattatta cggtcttcgc    46020

gcccgtacat gcattggacg attgcagggt caatggatct gagatcatcc agaggattgc    46080

cgcccttacc ttccgtttcg agttggagcc agcccctaaa tgagacgaca tagtcgactt    46140

gatgtgacaa tgccaagaga gagatttgct taacccgatt tttttgctca agcgtaagcc    46200

tattgaagct tgccggcatg acgtccgcgc cgaaagaata tcctacaagt aaaacattct    46260

gcacaccgaa atgcttggtg tagacatcga ttatgtgacc aagatcctta gcagtttcgc    46320

ttggggaccg ctccgaccag aaataccgaa gtgaactgac gccaatgaca ggaatccctt    46380

ccgtctgcag ataggtacca tcgatagatc tgctgcctcg cgcgtttcgg tgatgacggt    46440

gaaaacctct gacacatgca gctcccggag acggtcacag cttgtctgta agcggatgcc    46500

gggagcagac aagcccgtca gggcgcgtca gcgggtgttg gcgggtgtcg gggcgcagcc    46560

atgacccagt cacgtagcga tagcggagtg tatactggct taactatgcg gcatcagagc    46620

agattgtact gagagtgcac catatgcggt gtgaaatacc gcacagatgc gtaaggagaa    46680

aataccgcat caggcgctct ccgcttcct cgctcactga ctcgctgcgc tcggtcgttc    46740

ggctgcggcg agcggtatca gctcactcaa aggcggtaat acggttatcc acagaatcag    46800

gggataacgc aggaaagaac atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa    46860

aggccgcgtt gctggcgttt ttccataggc tccgcccccc tgacgagcat cacaaaaatc    46920

gacgctcaag tcagaggtgg cgaaacccga caggactata aagataccag gcgtttcccc    46980

ctggaagctc cctcgtgcgc tctcctgttc cgaccctgcc gcttaccgga tacctgtccg    47040

cctttctccc ttcgggaagc gtggcgcttt ctcatagctc acgctgtagg tatctcagtt    47100

cggtgtaggt cgttcgctcc aagctgggct gtgtgcacga accccccgtt cagcccgacc    47160

gctgcgcctt atccggtaac tatcgtcttg agtccaaccc ggtaagacac gacttatcgc    47220

cactggcagc agccactggt aacaggatta gcagagcgag gtatgtaggc ggtgctacag    47280

agttcttgaa gtggtggcct aactacggct acactagaag gacagtattt ggtatctgcg    47340

ctctgctgaa gccagttacc ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa    47400

ccaccgctgg tagcggtggt tttttgttt gcaagcagca gattacgcgc agaaaaaaag    47460

gatctcaaga agatcctttg atcttttcta cggggtctga cgctcagtgg aacgaaaact    47520
```

-continued

```
cacgttaagg gattttggtc atgagattat caaaaaggat cttcacctag atccttttaa    47580

attaaaaatg aagttttaaa tcaatctaaa gtatatatga gtaaacttgg tctgacagtt    47640

accaatgctt aatcagtgag gcacctatct cagcgatctg tctatttcgt tcatccatag    47700

ttgcctgact ccccgtcgtg tagataacta cgatacggga gggcttacca tctggcccca    47760

gtgctgcaat gataccgcga gacccacgct caccggctcc agatttatca gcaataaacc    47820

agccagccgg aagggccgag cgcagaagtg gtcctgcaac tttatccgcc tccatccagt    47880

ctattaattg ttgccgggaa gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg    47940

ttgttgccat tgctgcaggg gggggggggg ggggggactt ccattgttca ttccacggac    48000

aaaaacagag aaaggaaacg acagaggcca aaaagcctcg ctttcagcac ctgtcgtttc    48060

ctttcttttc agaggtatt ttaaataaaa acattaagtt atgacgaaga agaacgaaa    48120

cgccttaaac cggaaaattt tcataaatag cgaaaacccg cgaggtcgcc gccccgtagt    48180

cggatcaccg gaaaggaccc gtaaagtgat aatgattatc atctacatat cacaacgtgc    48240

gtggaggcca tcaaaccacg tcaaataatc aattatgacg caggtatcgt attaattgat    48300

ctgcatcaac ttaacgtaaa aacaacttca gacaatacaa atcagcgaca ctgaatacgg    48360

ggcaacctca tgtcccccc ccccccccc ctgcaggcat cgtggtgtca cgctcgtcgt    48420

ttggtatggc ttcattcagc tccggttccc aacgatcaag gcgagttaca tgatccccca    48480

tgttgtgcaa aaaagcggtt agctccttcg gtcctccgat cgttgtcaga gtaagttgg    48540

ccgcagtgtt atcactcatg gttatggcag cactgcataa ttctcttact gtcatgccat    48600

ccgtaagatg cttttctgtg actggtgagt actcaaccaa gtcattctga gaatagtgta    48660

tgcggcgacc gagttgctct tgcccggcgt caacacggga taataccgcg ccacatagca    48720

gaactttaaa agtgctcatc attggaaaac gttcttcggg gcgaaaactc tcaaggatct    48780

taccgctgtt gagatccagt tcgatgtaac ccactcgtgc acccaactga tcttcagcat    48840

cttttacttt caccagcgtt tctgggtgag caaaaacagg aaggcaaaat gccgcaaaaa    48900

agggaataag ggcgacacgg aaatgttgaa tactcatact cttccttttt caatattatt    48960

gaagcattta tcagggttat tgtctcatga gcggatacat atttgaatgt atttagaaaa    49020

ataaacaaat aggggttccg cgcacatttc cccgaaaagt gccacctgac gtctaagaaa    49080

ccattattat catgacatta acctataaaa ataggcgtat cacgaggccc tttcgtcttc    49140

aagaattggt cgacgatctt gctgcgttcg gatattttcg tggagttccc gccacagacc    49200

cggattgaag gcgagatcca gcaactcgcg ccagatcatc ctgtgacgga actttggcgc    49260

gtgatgactg gccaggacgt cggccgaaag agcgacaagc agatcacgct tttcgacagc    49320

gtcggatttg cgatcgagga ttttccggcg ctgcgctacg tccgcgaccg cgttgaggga    49380

tcaagccaca gcagcccact cgaccttcta gccgacccag acgagccaag ggatcttttt    49440

ggaatgctgc tccgtcgtca ggctttccga cgtttgggtg gttgaacaga agtcattatc    49500

gtacggaatg ccaagcactc ccgaggggaa ccctgtggtt ggcatgcaca tacaaatgga    49560

cgaacggata aaccttttca cgccctttta aatatccgtt attctaataa acgctctttt    49620

ctcttaggtt tacccgccaa tatatcctgt caaacactga tagtttaaac tgaaggcggg    49680

aaacgacaat ctgatcatga gcggagaatt aagggagtca cgttatgacc ccgccgatg    49740

acgcgggaca agccgtttta cgtttggaac tgacagaacc gcaacgttga aggagccact    49800
```

-continued

cagcaagctg gtacgattgt aatacgactc actatagggc gaattgagcg ctgtttaaac   49860

gctcttcaac tggaagagcg gttacccgga ccgaagcttg catgcctgca g   49911

<210> SEQ ID NO 32
<211> LENGTH: 36909
<212> TYPE: DNA
<213> ORGANISM: artificial
<220> FEATURE:
<223> OTHER INFORMATION: vector

<400> SEQUENCE: 32

tctagagctc gttcctcgag gcctcgaggc ctcgaggaac ggtacctgcg gggaagctta   60

caataatgtg tgttgttaag tcttgttgcc tgtcatcgtc tgactgactt tcgtcataaa   120

tcccggcctc cgtaacccag ctttgggcaa gctcacggat ttgatccggc ggaacgggaa   180

tatcgagatg ccgggctgaa cgctgcagtt ccagctttcc ctttcgggac aggtactcca   240

gctgattgat tatctgctga agggtcttgg ttccacctcc tggcacaatg cgaatgatta   300

cttgagcgcg atcgggcatc caattttctc ccgtcaggtg cgtggtcaag tgctacaagg   360

cacctttcag taacgagcga ccgtcgatcc gtcgccggga tacggacaaa atggagcgca   420

gtagtccatc gagggcggcg aaagcctcgc caaaagcaat acgttcatct cgcacagcct   480

ccagatccga tcgagggtct tcggcgtagg cagatagaag catggataca ttgcttgaga   540

gtattccgat ggactgaagt atggcttcca tcttttctcg tgtgtctgca tctatttcga   600

gaaagccccc gatgcggcgc accgcaacgc gaattgccat actatccgaa agtcccagca   660

ggcgcgcttg ataggaaaag gtttcatact cggccgatcg cagacgggca ctcacgacct   720

tgaacccttc aactttcagg gatcgatgct ggttgatggt agtctcactc gacgtggctc   780

tggtgtgttt tgacatagct tcctccaaag aaagcggaag gtctggatac tccagcacga   840

aatgtgcccg ggtagacgga tggaagtcta gccctgctca atatgaaatc aacagtacat   900

ttacagtcaa tactgaatat acttgctaca tttgcaattg tcttataacg aatgtgaaat   960

aaaaatagtg taacaacgct tttactcatc gataatcaca aaaacattta tacgaacaaa   1020

aatacaaatg cactccggtt tcacaggata ggcgggatca gaatatgcaa cttttgacgt   1080

tttgttcttt caaagggggt gctggcaaaa ccaccgcact catgggcctt gcgctgctt   1140

tggcaaatga cggtaaacga gtggccctct ttgatgccga cgaaaaccgg cctctgacgc   1200

gatggagaga aaacgcctta caaagcagta ctgggatcct cgctgtgaag tctattccgc   1260

cgacgaaatg cccttcttg aagcagccta tgaaaatgcc gagctcgaag gatttgatta   1320

tgcgttggcc gatacgcgtg cggctcgag cgagctcaac aacacaatca tcgctagctc   1380

aaacctgctt ctgatcccca ccatgctaac gccgctcgac atcgatgagg cactatctac   1440

ctaccgctac gtcatcgagc tgctgttgag tgaaaatttg gcaattccta cagctgtttt   1500

gcgccaacgc gtcccggtcg gccgattgac aacatcgcaa cgcaggatgt cagagacgct   1560

agagagcctt ccagttgtac cgtctcccat gcatgaaaga gatgcatttg ccgcgatgaa   1620

agaacgcggc atgttgcatc ttacattact aaacacggga actgatccga cgatgcgcct   1680

catagagagg aatcttcgga ttgcgatgga ggaagtcgtg gtcatttcga aactgatcag   1740

caaaatcttg gaggcttgaa gatggcaatt cgcaagcccg cattgtcggt cggcgaagca   1800

cggcggcttg ctggtgctcg acccgagatc caccatccca acccgacact tgttccccag   1860

-continued

```
aagctggacc tccagcactt gcctgaaaaa gccgacgaga aagaccagca acgtgagcct    1920

ctcgtcgccg atcacattta cagtcccgat cgacaactta agctaactgt ggatgccctt    1980

agtccacctc cgtccccgaa aaagctccag gttttctttt cagcgcgacc gcccgcgcct    2040

caagtgtcga aaacatatga caacctcgtt cggcaataca gtccctcgaa gtcgctacaa    2100

atgattttaa ggcgcgcgtt ggacgatttc gaaagcatgc tggcagatgg atcatttcgc    2160

gtggccccga aagttatcc gatcccttca actacagaaa aatccgttct cgttcagacc     2220

tcacgcatgt cccggttgc gttgctcgag gtcgctcgaa gtcattttga tccgttgggg     2280

ttggagaccg ctcgagcttt cggccacaag ctggctaccg ccgcgctcgc gtcattcttt    2340

gctggagaga agccatcgag caattggtga agagggacct atcggaaccc ctcaccaaat    2400

attgagtgta ggtttgaggc cgctggccgc gtcctcagtc accttttgag ccagataatt    2460

aagagccaaa tgcaattggc tcaggctgcc atcgtccccc cgtgcgaaac ctgcacgtcc    2520

gcgtcaaaga aataaccggc acctcttgct gtttttatca gttgagggct tgacggatcc    2580

gcctcaagtt tgcggcgcag ccgcaaaatg agaacatcta tactcctgtc gtaaacctcc    2640

tcgtcgcgta ctcgactggc aatgagaagt tgctcgcgcg atagaacgtc gcggggtttc    2700

tctaaaaacg cgaggagaag attgaactca cctgccgtaa gtttcacctc accgccagct    2760

tcggacatca agcgacgttg cctgagatta agtgtccagt cagtaaaaca aaaagaccgt    2820

cggtctttgg agcggacaac gttggggcgc acgcgcaagg caacccgaat gcgtgcaaga    2880

aactctctcg tactaaacgg cttagcgata aaatcacttg ctcctagctc gagtgcaaca    2940

actttatccg tctcctcaag gcggtcgcca ctgataatta tgattggaat atcagacttt    3000

gccgccagat ttcgaacgat ctcaagccca tcttcacgac ctaaatttag atcaacaacc    3060

acgacatcga ccgtcgcgga agagagtact ctagtgaact gggtgctgtc ggctaccgcg    3120

gtcactttga aggcgtggat cgtaaggtat tcgataataa gatgccgcat agcgacatcg    3180

tcatcgataa gaagaacgtg tttcaacggc tcacctttca atctaaaatc tgaacccttg    3240

ttcacagcgc ttgagaaatt ttcacgtgaa ggatgtacaa tcatctccag ctaaatgggc    3300

agttcgtcag aattgcggct gaccgcggat gacgaaaatg cgaaccaagt atttcaattt    3360

tatgacaaaa gttctcaatc gttgttacaa gtgaaacgct tcgaggttac agctactatt    3420

gattaaggag atcgcctatg gtctcgcccc ggcgtcgtgc gtccgccgcg agccagatct    3480

cgcctacttc ataaacgtcc tcataggcac ggaatggaat gatgacatcg atcgccgtag    3540

agagcatgtc aatcagtgtg cgatcttcca agctagcacc ttgggcgcta cttttgacaa    3600

gggaaaacag tttcttgaat ccttggattg gattcgcgcc gtgtattgtt gaaatcgatc    3660

ccggatgtcc cgagacgact tcactcagat aagcccatgc tgcatcgtcg cgcatctcgc    3720

caagcaatat ccggtccggc cgcatacgca gacttgcttg agcaagtgc tcggcgctca     3780

cagcacccag cccagcaccg ttcttggagt agagtagtct aacatgatta tcgtgtggaa    3840

tgacgagttc gagcgtatct tctatggtga ttagccttc ctgggggggg atggcgctga     3900

tcaaggtctt gctcattgtt gtcttgccgc ttccggtagg ccacatagc aacatcgtca     3960

gtcggctgac gacgcatgcg tgcagaaacg cttccaaatc cccgttgtca aaatgctgaa    4020

ggatagcttc atcatcctga ttttggcgtt tccttcgtgt ctgccactgg ttccacctcg    4080

aagcatcata acgggaggag acttctttaa gaccagaaac acgcgagctt ggccgtcgaa    4140
```

-continued

```
tggtcaagct gacggtgccc gagggaacgg tcggcggcag acagatttgt agtcgttcac    4200

caccaggaag ttcagtggcg cagagggggt tacgtggtcc gacatcctgc tttctcagcg    4260

cgcccgctaa aatagcgata tcttcaagat catcataaga gacgggcaaa ggcatcttgg    4320

taaaaatgcc ggcttggcgc acaaatgcct ctccaggtcg attgatcgca atttcttcag    4380

tcttcgggtc atcgagccat tccaaaatcg gcttcagaag aaagcgtagt tgcggatcca    4440

cttccattta caatgtatcc tatctctaag cggaaatttg aattcattaa gagcggcggt    4500

tcctcccccg cgtggcgccg ccagtcaggc ggagctggta aacaccaaag aaatcgaggt    4560

cccgtgctac gaaaatggaa acggtgtcac cctgattctt cttcagggtt ggcggtatgt    4620

tgatggttgc cttaagggct gtctcagttg tctgctcacc gttattttga aagctgttga    4680

agctcatccc gccacccgag ctgccggcgt aggtgctagc tgcctggaag cgcgccttgaa    4740

caacactcaa gagcatagct ccgctaaaac gctgccagaa gtggctgtcg accgagcccg    4800

gcaatcctga cgaccgagt tcgtccgcgc ttggcgatgt taacgagatc atcgcatggt    4860

caggtgtctc ggcgcgatcc cacaacacaa aaacgcgccc atctccctgt tgcaagccac    4920

gctgtatttc gccaacaacg gtggtgccac gatcaagaag cacgatattg ttcgttgttc    4980

cacgaatatc ctgaggcaag acacacttta catagcctgc caaatttgtg tcgattgcgg    5040

tttgcaagat gcacggaatt attgtcccctt gcgttaccat aaaatcgggg tgcggcaaga    5100

gcgtggcgct gctgggctgc agctcggtgg gtttcatacg tatcgacaaa tcgttctcgc    5160

cggacacttc gccattcggc aaggagttgt cgtcacgctt gccttcttgt cttcggcccg    5220

tgtcgccctg aatggcgcgt ttgctgaccc cttgatcgcc gctgctatat gcaaaaatcg    5280

gtgtttcttc cggccgtggc tcatgccgct ccggttcgcc cctcggcggt agaggagcag    5340

caggctgaac agcctcttga accgctggag gatccggcgg cacctcaatc ggagctggat    5400

gaaatggctt ggtgtttgtt gcgatcaaag ttgacggcga tgcgttctca ttcaccttct    5460

tttggcgccc acctagccaa atgaggctta atgataacgc gagaacgaca cctccgacga    5520

tcaatttctg agaccccgaa agacgccggc gatgtttgtc ggagaccagg atccagatg    5580

catcaacctc atgtgccgct tgctgactat cgttattcat cccttcgccc ccttcaggac    5640

gcgtttcaca tcgggcctca ccgtgccgt ttgcggcctt tggccaacgg atcgtaagc    5700

ggtgttccag atacatagta ctgtgtggcc atccctcaga cgccaacctc gggaaaccga    5760

agaaatctcg acatcgctcc ctttaactga atagttggca acagcttcct tgccatcagg    5820

attgatggtg tagatggagg gtatgcgtac attgcccgga aagtggaata ccgtcgtaaa    5880

tccattgtcg aagacttcga gtggcaacag cgaacgatcg ccttgggcga cgtagtgcca    5940

attactgtcc gccgcaccaa gggctgtgac aggctgatcc aataaattct cagctttccg    6000

ttgatattgt gcttccgcgt gtagtctgtc cacaacagcc ttctgttgtg cctcccttcg    6060

ccgagccgcc gcatcgtcgg cggggtaggc gaattggacg ctgtaataga gatcgggctg    6120

ctctttatcg aggtgggaca gagtcttgga acttatactg aaaacataac ggcgcatccc    6180

ggagtcgctt gcggttagca cgattactgg ctgaggcgtg aggacctggc ttgccttgaa    6240

aaatagataa tttccccgcg gtagggctgc tagatctttg ctatttgaaa cggcaaccgc    6300

tgtcaccgtt tcgttcgtgg cgaatgttac gaccaaagta gctccaaccg ccgtcgagag    6360

gcgcaccact tgatcgggat tgtaagccaa ataacgcatg cgcggatcta gcttgcccgc    6420
```

-continued

```
cattggagtg tcttcagcct ccgcaccagt cgcagcggca aataaacatg ctaaaatgaa    6480

aagtgctttt ctgatcatgg ttcgctgtgg cctacgtttg aaacggtatc ttccgatgtc    6540

tgataggagg tgacaaccag acctgccggg ttggttagtc tcaatctgcc gggcaagctg    6600

gtcacctttt cgtagcgaac tgtcgcggtc cacgtactca ccacaggcat tttgccgtca    6660

acgacgaggg tcctttata gcgaatttgc tgcgtgcttg gagttacatc atttgaagcg    6720

atgtgctcga cctccaccct gccgcgtttg ccaagaatga cttgaggcga actgggattg    6780

ggatagttga agaattgctg gtaatcctgg cgcactgttg gggcactgaa gttcgatacc    6840

aggtcgtagg cgtactgagc ggtgtcggca tcataactct cgcgcaggcg aacgtactcc    6900

cacaatgagg cgttaacgac ggcctcctct tgagttgcag gcaatcgcga gacagacacc    6960

tcgctgtcaa cggtgccgtc cggccgtatc catagatata cgggcacaag cctgctcaac    7020

ggcaccattg tggctatagc gaacgcttga gcaacatttc ccaaaatcgc gatagctgcg    7080

acagctgcaa tgagtttgga gagacgtcgc gccgatttcg ctcgcgcggt ttgaaaggct    7140

tctacttcct tatagtgctc ggcaaggctt cgcgcgcca ctagcatggc atattcaggc    7200

cccgtcatag cgtccacccg aattgccgag ctgaagatct gacggagtag gctgccatcg    7260

ccccacattc agcgggaaga tcgggccttt gcagctcgct aatgtgtcgt ttgtctggca    7320

gccgctcaaa gcgacaacta ggcacagcag gcaatacttc atagaattct ccattgaggc    7380

gaatttttgc gcgacctagc ctcgctcaac ctgagcgaag cgacggtaca agctgctggc    7440

agattgggtt gcgccgctcc agtaactgcc tccaatgttg ccggcgatcg ccggcaaagc    7500

gacaatgagc gcatccctg tcagaaaaaa catatcgagt tcgtaaagac caatgatctt    7560

ggccgcggtc gtaccggcga aggtgattac accaagcata agggtgagcg cagtcgcttc    7620

ggttaggatg acgatcgttg ccacgaggtt taagaggaga agcaagagac cgtaggtgat    7680

aagttgcccg atccacttag ctgcgatgtc ccgcgtgcga tcaaaaatat atccgacgag    7740

gatcagaggc ccgatcgcga gaagcacttt cgtgagaatt ccaacggcgt cgtaaactcc    7800

gaaggcagac cagagcgtgc cgtaaaggac ccactgtgcc ccttggaaag caaggatgtc    7860

ctggtcgttc atcggaccga tttcggatgc gattttctga aaaacggcct gggtcacggc    7920

gaacattgta tccaactgtg ccggaacagt ctgcagaggc aagccggtta cactaaactg    7980

ctgaacaaag tttgggaccg tcttttcgaa gatggaaacc acatagtctt ggtagttagc    8040

ctgcccaaca attagagcaa caacgatggt gaccgtgatc acccgagtga taccgctacg    8100

ggtatcgact tcgccgcgta tgactaaaat accctgaaca ataatccaaa gagtgacaca    8160

ggcgatcaat ggcgcactca ccgcctcctg gatagtctca agcatcgagt ccaagcctgt    8220

cgtgaaggct acatcgaaga tcgtatgaat ggccgtaaac ggcgccggaa tcgtgaaatt    8280

catcgattgg acctgaactt gactggtttg tcgcataatg ttggataaaa tgagctcgca    8340

ttcggcgagg atgcgggcgg atgaacaaat cgcccagcct tagggagggg caccaaagat    8400

gacagcggtc ttttgatgct ccttgcgttg agcggccgcc tcttccgcct cgtgaaggcc    8460

ggcctgcgcg gtagtcatcg ttaataggct tgtcgcctgt acattttgaa tcattgcgtc    8520

atggatctgc ttgagaagca aaccattggt cacggttgcc tgcatgatat tgcgagatcg    8580

ggaaagctga gcagacgtat cagcattcgc cgtcaagcgt ttgtccatcg tttccagatt    8640

gtcagccgca atgccagcgc tgtttgcgga accggtgatc tgcgatcgca acaggtccgc    8700
```

```
ttcagcatca ctacccacga ctgcacgatc tgtatcgctg gtgatcgcac gtgccgtggt    8760

cgacattggc attcgcggcg aaaacatttc attgtctagg tccttcgtcg aaggatactg    8820

attttttctgg ttgagcgaag tcagtagtcc agtaacgccg taggccgacg tcaacatcgt    8880

aaccatcgct atagtctgag tgagattctc cgcagtcgcg agcgcagtcg cgagcgtctc    8940

agcctccgtt gccgggtcgc taacaacaaa ctgcgcccgc gcgggctgaa tatatagaaa    9000

gctgcaggtc aaaactgttg caataagttg cgtcgtcttc atcgtttcct accttatcaa    9060

tcttctgcct cgtggtgacg ggccatgaat tcgctgagcc agccagatga gttgccttct    9120

tgtgcctcgc gtagtcgagt tgcaaagcgc accgtgttgg cacgccccga aagcacggcg    9180

acatattcac gcatatcccg cagatcaaat tcgcagatga cgcttccact ttctcgttta    9240

agaagaaact tacggctgcc gaccgtcatg tcttcacgga tcgcctgaaa ttccttttcg    9300

gtacatttca gtccatcgac ataagccgat cgatctgcgg ttggtgatgg atagaaaatc    9360

ttcgtcatac attgcgcaac caagctggct cctagcggcg attccagaac atgctctggt    9420

tgctgcgttg ccagtattag catcccgttg ttttttcgaa cggtcaggag gaatttgtcg    9480

acgacagtcg aaaatttagg gtttaacaaa taggcgcgaa actcatcgca gctcatcaca    9540

aaacggcggc cgtcgatcat ggctccaatc cgatgcagga gatatgctgc agcgggagcg    9600

catacttcct cgtattcgag aagatgcgtc atgtcgaagc cggtaatcga cggatctaac    9660

tttacttcgt caacttcgcc gtcaaatgcc cagccaagcg catggccccg gcaccagcgt    9720

tggagccgcg ctcctgcgcc ttcggcgggc ccatgcaaca aaaattcacg taaccccgcg    9780

attgaacgca tttgtggatc aaacgagagc tgacgatgga taccacggac cagacggcgg    9840

ttctcttccg gagaaatccc accccgacca tcactctcga tgagagccac gatccattcg    9900

cgcagaaaat cgtgtgaggc tgctgtgttt tctaggccac gcaacggcgc caacccgctg    9960

ggtgtgcctc tgtgaagtgc caaatatgtt cctcctgtgg cgcgaaccag caattcgcca    10020

ccccggtcct tgtcaaagaa cacgaccgta cctgcacggt cgaccatgct ctgttcgagc    10080

atggctagaa caaacatcat gagcgtcgtc ttacccctcc cgataggccc gaatattgcc    10140

gtcatgccaa catcgtgctc atgcgggata tagtcgaaag gcgttccgcc attggtacga    10200

aatcgggcaa tcgcgttgcc ccagtggcct gagctggcgc cctctggaaa gttttcgaaa    10260

gagacaaacc ctgcgaaatt gcgtgaagtg attgcgccag ggcgtgtgcg ccacttaaaa    10320

ttccccggca attgggacca ataggccgct tccataccaa taccttcttg gacaaccacg    10380

gcacctgcat ccgccattcg tgtccgagcc gcgcgcgcccc tgtccccaag actattgaga    10440

tcgtctgcat agacgcaaag gctcaaatga tgtgagccca taacgaattc gttgctcgca    10500

agtgcgtcct cagcctcgga taatttgccg atttgagtca cggctttatc gccggaactc    10560

agcatctggc tcgatttgag gctaagtttc gcgtgcgctt gcgggcgagt caggaacgaa    10620

aaactctgcg tgagaacaag tggaaaatcg agggatagca gcgcgttgag catgcccggc    10680

cgtgtttttg cagggtattc gcgaaacgaa tagatggatc caacgtaact gtcttttggc    10740

gttctgatct cgagtcctcg cttgccgcaa atgactctgt cggtataaat cgaagcgccg    10800

agtgagccgc tgacgaccgg aaccggtgtg aaccgaccag tcatgatcaa ccgtagcgct    10860

tcgccaattt cggtgaagag cacaccctgc ttctcgcgga tgccaagacg atgcaggcca    10920

tacgctttaa gagagccagc gacaacatgc caaagatctt ccatgttcct gatctggccc    10980
```

-continued

```
gtgagatcgt tttccctttt tccgcttagc ttggtgaacc tcctctttac cttccctaaa   11040

gccgcctgtg ggtagacaat caacgtaagg aagtgttcat tgcggaggag ttggccggag   11100

agcacgcgct gttcaaaagc ttcgttcagg ctagcggcga aaacactacg gaagtgtcgc   11160

ggcgccgatg atggcacgtc ggcatgacgt acgaggtgag catatattga cacatgatca   11220

tcagcgatat tgcgcaacag cgtgttgaac gcacgacaac gcgcattgcg catttcagtt   11280

tcctcaagct cgaatgcaac gccatcaatt ctcgcaatgg tcatgatcga tccgtcttca   11340

agaaggacga tatggtcgct gaggtggcca atataaggga gatagatctc accggatctt   11400

tcggtcgttc cactcgcgcc gagcatcaca ccattcctct ccctcgtggg ggaaccctaa   11460

ttggatttgg gctaacagta gcgccccccc aaactgcact atcaatgctt cttcccgcgg   11520

tccgcaaaaa tagcaggacg acgctcgccg cattgtagtc tcgctccacg atgagccggg   11580

ctgcaaacca taacggcacg agaacgactt cgtagagcgg gttctgaacg ataacgatga   11640

caaagccggc gaacatcatg aataaccctg ccaatgtcag tggcaccca agaaacaatg   11700

cgggccgtgt ggctgcgagg taaagggtcg attcttccaa acgatcagcc atcaactacc   11760

gccagtgagc gtttggccga ggaagctcgc cccaaacatg ataacaatgc cgccgacgac   11820

gccggcaacc agcccaagcg aagcccgccc gaacatccag gagatcccga tagcgacaat   11880

gccgagaaca gcgagtgact ggccgaacgg accaaggata aacgtgcata tattgttaac   11940

cattgtggcg gggtcagtgc cgccacccgc agattgcgct gcggcgggtc cggatgagga   12000

aatgctccat gcaattgcac cgcacaagct tggggcgcag ctcgatatca gcgcgcatcat   12060

cgcattcgag agcgagaggc gatttagatg taaacggtat ctctcaaagc atcgcatcaa   12120

tgcgcacctc cttagtataa gtcgaataag acttgattgt cgtctgcgga tttgccgttg   12180

tcctggtgtg gcggtggcgg agcgattaaa ccgccagcgc catcctcctg cgagcggcgc   12240

tgatatgacc cccaaacatc ccacgtctct tcggatttta cgcctcgtg atcgtctttt   12300

ggaggctcga ttaacgcggg caccagcgat tgagcagctg tttcaacttt tcgcacgtag   12360

ccgtttgcaa aaccgccgat gaaattaccg gtgttgtaag cggagatcgc ccgacgaagc   12420

gcaaattgct tctcgtcaat cgtttcgccg cctgcataac gactttttcag catgtttgca   12480

gcggcagata atgatgtgca cgcctggagc gcaccgtcag gtgtcagacc gagcatagaa   12540

aaatttcgag agtttatttg catgaggcca acatccagcg aatgccgtgc atcgagacgg   12600

tgcctgacga cttgggttgc ttggctgtga tcttgccagt gaagcgtttc gccggtcgtg   12660

ttgtcatgaa tcgctaaagg atcaaagcga ctctccacct tagctatcgc cgcaagcgta   12720

gatgtcgcaa ctgatggggc acacttgcga gcaacatggt caaactcagc agatgagagt   12780

ggcgtggcaa ggctcgacga acagaaggag accatcaagg caagagaaag cgaccccgat   12840

ctcttaagca taccttatct ccttagctcg caactaacac cgcctctccc gttggaagaa   12900

gtgcgttgtt ttatgttgaa gattatcggg agggtcggtt actcgaaaat tttcaattgc   12960

ttctttatga tttcaattga agcgagaaac ctcgcccggc gtcttggaac gcaacatgga   13020

ccgagaaccg cgcatccatg actaagcaac cggatcgacc tattcaggcc gcagttggtc   13080

aggtcaggct cagaacgaaa atgctcggcg aggttacgct gtctgtaaac ccattcgatg   13140

aacgggaagc ttccttccga ttgctcttgg caggaatatt ggcccatgcc tgcttgcgct   13200

ttgcaaatgc tcttatcgcg ttggtatcat atgccttgtc cgccagcaga aacgcactct   13260
```

-continued

```
aagcgattat ttgtaaaaat gtttcggtca tgcggcggtc atgggcttga cccgctgtca   13320

gcgcaagacg gatcggtcaa ccgtcggcat cgacaacagc gtgaatcttg gtggtcaaac   13380

cgccacggga acgtcccata cagccatcgt cttgatcccg ctgtttcccg tcgccgcatg   13440

ttggtggacg cggacacagg aactgtcaat catgacgaca ttctatcgaa agccttggaa   13500

atcacactca gaatatgatc ccagacgtct gcctcacgcc atcgtacaaa gcgattgtag   13560

caggttgtac aggaaccgta tcgatcagga acgtctgccc agggcgggcc cgtccggaag   13620

cgccacaaga tgacattgat cacccgcgtc aacgcgcggc acgcgacgcg gcttatttgg   13680

gaacaaagga ctgaacaaca gtccattcga aatcggtgac atcaaagcgg ggacgggtta   13740

tcagtggcct ccaagtcaag cctcaatgaa tcaaaatcag accgatttgc aaacctgatt   13800

tatgagtgtg cggcctaaat gatgaaatcg tccttctaga tcgcctccgt ggtgtagcaa   13860

cacctcgcag tatcgccgtg ctgaccttgg ccagggaatt gactggcaag ggtgctttca   13920

catgaccgct cttttggccg cgatagatga tttcgttgct gctttgggca cgtagaagga   13980

gagaagtcat atcggagaaa ttcctcctgg cgcgagagcc tgctctatcg cgacggcatc   14040

ccactgtcgg gaacagaccg gatcattcac gaggcgaaag tcgtcaacac atgcgttata   14100

ggcatcttcc cttgaaggat gatcttgttg ctgccaatct ggaggtgcgg cagccgcagg   14160

cagatgcgat ctcagcgcaa cttgcggcaa aacatctcac tcacctgaaa accactagcg   14220

agtctcgcga tcagacgaag gcctttact taacgacaca atatccgatg tctgcatcac   14280

aggcgtcgct atcccagtca atactaaagc ggtgcaggaa ctaaagatta ctgatgactt   14340

aggcgtgcca cgaggcctga gacgacgcgc gtagacagtt ttttgaaatc attatcaaag   14400

tgatggcctc cgctgaagcc tatcacctct gcgccggtct gtcggagaga tgggcaagca   14460

ttattacggt cttcgcgccc gtacatgcat tggacgattg cagggtcaat ggatctgaga   14520

tcatccagag gattgccgcc cttaccttcc gtttcgagtt ggagccagcc cctaaatgag   14580

acgacatagt cgacttgatg tgacaatgcc aagagagaga tttgcttaac ccgatttttt   14640

tgctcaagcg taagcctatt gaagcttgcc ggcatgacgt ccgcgccgaa agaatatcct   14700

acaagtaaaa cattctgcac accgaaatgc ttggtgtaga catcgattat gtgaccaaga   14760

tccttagcag tttcgcttgg ggaccgctcc gaccagaaat accgaagtga actgacgcca   14820

atgacaggaa tcccttccgt ctgcagatag gtaccatcga tagatctgct gcctcgcgcg   14880

tttcggtgat gacggtgaaa acctctgaca catgcagctc ccggagacgg tcacagcttg   14940

tctgtaagcg gatgccggga gcagacaagc ccgtcagggc gcgtcagcgg gtgttggcgg   15000

gtgtcggggc gcagccatga cccagtcacg tagcgatagc ggagtgtata ctggcttaac   15060

tatgcggcat cagagcagat tgtactgaga gtgcaccata tgcggtgtga aataccgcac   15120

agatgcgtaa ggagaaaata ccgcatcagg cgctcttccg cttcctcgct cactgactcg   15180

ctgcgctcgg tcgttcggct gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg   15240

ttatccacag aatcagggga taacgcagga agaacatgt gagcaaaagg ccagcaaaag   15300

gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg cccccctgac   15360

gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg actataaaga   15420

taccaggcgt ttccccctgg aagctccctc gtgcgctctc ctgttccgac cctgccgctt   15480

accggatacc tgtccgcctt tctcccttcg ggaagcgtgg cgctttctca tagctcacgc   15540
```

-continued

```
tgtaggtatc tcagttcggt gtaggtcgtt cgctccaagc tgggctgtgt gcacgaaccc   15600

cccgttcagc ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc caacccggta   15660

agacacgact tatcgccact ggcagcagcc actggtaaca ggattagcag agcgaggtat   15720

gtaggcggtg ctacagagtt cttgaagtgg tggcctaact acggctacac tagaaggaca   15780

gtatttggta tctgcgctct gctgaagcca gttaccttcg aaaaagagt  tggtagctct   15840

tgatccggca aacaaccac  cgctggtagc ggtggttttt ttgtttgcaa gcagcagatt   15900

acgcgcagaa aaaaggatc  tcaagaagat cctttgatct tttctacggg gtctgacgct   15960

cagtggaacg aaaactcacg ttaagggatt ttggtcatga gattatcaaa aaggatcttc   16020

acctagatcc ttttaaatta aaaatgaagt tttaaatcaa tctaaagtat atatgagtaa   16080

acttggtctg acagttacca atgcttaatc agtgaggcac ctatctcagc gatctgtcta   16140

tttcgttcat ccatagttgc ctgactcccc gtcgtgtaga taactacgat acgggagggc   16200

ttaccatctg gccccagtgc tgcaatgata ccgcgagacc cacgctcacc ggctccagat   16260

ttatcagcaa taaaccagcc agccggaagg gccgagcgca gaagtggtcc tgcaacttta   16320

tccgcctcca tccagtctat taattgttgc cgggaagcta gagtaagtag ttcgccagtt   16380

aatagtttgc gcaacgttgt tgccattgct gcaggggggg ggggggggg  gttccattgt   16440

tcattccacg acaaaaaca  gagaaaggaa acgacagagg ccaaaaagct cgctttcagc   16500

acctgtcgtt tcctttcttt tcagagggta ttttaaataa aaacattaag ttatgacgaa   16560

gaagaacgga aacgcctaa  accggaaaat tttcataaat agcgaaaacc cgcgaggtcg   16620

ccgccccgta acctgtcgga tcaccggaaa ggacccgtaa agtgataatg attatcatct   16680

acatatcaca acgtgcgtgg aggccatcaa accacgtcaa ataatcaatt atgacgcagg   16740

tatcgtatta attgatctgc atcaacttaa cgtaaaaaca acttcagaca atacaaatca   16800

gcgacactga atacggggca acctcatgtc ccccccccc  ccccccctgc aggcatcgtg   16860

gtgtcacgct cgtcgtttgg tatggcttca ttcagctccg gttcccaacg atcaaggcga   16920

gttacatgat cccccatgtt gtgcaaaaaa gcggttagct ccttcggtcc tccgatcgtt   16980

gtcagaagta agttggccgc agtgttatca ctcatggtta tggcagcact gcataattct   17040

cttactgtca tgccatccgt aagatgcttt tctgtgactg gtgagtactc aaccaagtca   17100

ttctgagaat agtgtatgcg cgaccgagt  tgctcttgcc cggcgtcaac acgggataat   17160

accgcgccac atagcagaac tttaaaagtg ctcatcattg gaaaacgttc ttcggggcga   17220

aaactctcaa ggatcttacc gctgttgaga tccagttcga tgtaacccac tcgtgcaccc   17280

aactgatctt cagcatcttt tactttcacc agcgtttctg ggtgagcaaa aacaggaagg   17340

caaaatgccg caaaaaaggg aataagggc  acacggaaat gttgaatact catactcttc   17400

ctttttcaat attattgaag catttatcag ggttattgtc tcatgagcgg atacatattt   17460

gaatgtattt agaaaaataa acaaataggg gttccgcgca catttccccg aaaagtgcca   17520

cctgacgtct aagaaaccat tattatcatg acattaacct ataaaaatag gcgtatcacg   17580

aggccctttc gtcttcaaga attcggagct tttgccattc tcaccggatt cagtcgtcac   17640

tcatggtgat ttctcacttg ataaccttat ttttgacgag gggaaattaa taggttgtat   17700

tgatgttgga cgagtcggaa tcgcagaccg ataccaggat cttgccatcc tatggaactg   17760

cctcggtgag ttttctcctt cattacagaa acggcttttt caaaaatatg gtattgataa   17820
```

```
tcctgatatg aataaattgc agtttcattt gatgctcgat gagttttct aatcagaatt    17880

ggttaattgg ttgtaacact ggcagagcat tacgctgact tgacgggacg gcggctttgt    17940

tgaataaatc gaacttttgc tgagttgaag gatcagatca cgcatcttcc cgacaacgca    18000

gaccgttccg tggcaaagca aaagttcaaa atcaccaact ggtccaccta caacaaagct    18060

ctcatcaacc gtggctccct cactttctgg ctggatgatg gggcgattca ggcctggtat    18120

gagtcagcaa caccttcttc acgaggcaga cctcagcgcc agaaggccgc cagagaggcc    18180

gagcgcggcc gtgaggcttg gacgctaggg cagggcatga aaaagcccgt agcgggctgc    18240

tacgggcgtc tgacgcggtg gaaaggggga ggggatgttg tctacatggc tctgctgtag    18300

tgagtgggtt gcgctccggc agcggtcctg atcaatcgtc acctttctc ggtccttcaa     18360

cgttcctgac aacgagcctc cttttcgcca atccatcgac aatcaccgcg agtccctgct    18420

cgaacgctgc gtccggaccg gcttcgtcga aggcgtctat cgcggcccgc aacagcggcg    18480

agagcggagc ctgttcaacg gtgccgccgc gctcgccggc atcgctgtcg ccggcctgct    18540

cctcaagcac ggccccaaca gtgaagtagc tgattgtcat cagcgcattg acggcgtccc    18600

cggccgaaaa acccgcctcg cagaggaagc gaagctgcgc gtcggccgtt ccatctgcg     18660

gtgcgcccgg tcgcgtgccg gcatggatgc gcgcgccatc gcggtaggcg agcagcgcct    18720

gcctgaagct gcgggcattc ccgatcagaa atgagcgcca gtcgtcgtcg gctctcggca    18780

ccgaatgcgt atgattctcc gccagcatgg cttcggccag tgcgtcgagc agcgcccgct    18840

tgttcctgaa gtgccagtaa agcgccggct gctgaacccc caaccgttcc gccagtttgc    18900

gtgtcgtcag accgtctacg ccgacctcgt tcaacaggtc cagggcggca cggatcactg    18960

tattcggctg caactttgtc atgcttgaca ctttatcact gataaacata atatgtccac    19020

caacttatca gtgataaaga atccgcgcgt tcaatcggac cagcggaggc tggtccggag    19080

gccagacgtg aaacccaaca tacccctgat cgtaattctg agcactgtcg cgctcgacgc    19140

tgtcggcatc ggcctgatta tgccggtgct gccgggcctc ctgcgcgatc tggttcactc    19200

gaacgacgtc accgcccact atggcattct gctggcgctg tatgcgttgg tgcaatttgc    19260

ctgcgcacct gtgctgggcg cgctgtcgga tcgtttcggg cggcggccaa tcttgctcgt    19320

ctcgctggcc ggcgccactg tcgactacgc catcatggcg acagcgcctt cctttgggt     19380

tctctatatc gggcggatcg tggccggcat caccgggggcg actggggcgg tagccggcgc    19440

ttatattgcc gatatcactg atggcgatga gcgcgcgcgg cacttcggct tcatgagcgc    19500

ctgtttcggg ttcgggatgg tcgcgggacc tgtgctcggt gggctgatgg gcggtttctc    19560

cccccacgct ccgttcttcg ccgcggcagc cttgaacggc ctcaatttcc tgacgggctg    19620

tttcctttg ccggagtcgc acaaaggcga acgccggccg ttacgccggg aggctctcaa     19680

cccgctcgct tcgttccggt gggcccgggg catgaccgtc gtcgccgccc tgatggcggt    19740

cttcttcatc atgcaacttg tcggacaggt gccggccgcg ctttgggtca ttttcggcga    19800

ggatcgcttt cactgggacg cgaccacgat cggcatttcg cttgccgcat ttggcattct    19860

gcattcactc gcccaggcaa tgatcaccgg ccctgtagcc gccggctcg gcgaaaggcg     19920

ggcactcatg ctcggaatga ttgccgacgg cacaggctac atcctgcttg ccttcgcgac    19980

acggggatgg atggcgttcc cgatcatggt cctgcttgct tcgggtggca tcggaatgcc    20040

ggcgctgcaa gcaatgttgt ccaggcaggt ggatgaggaa cgtcagggc agctgcaagg     20100
```

-continued

```
ctcactggcg gcgctcacca gcctgacctc gatcgtcgga cccctcctct tcacggcgat    20160

ctatgcggct tctataacaa cgtggaacgg gtgggcatgg attgcaggcg ctgccctcta    20220

cttgctctgc ctgccggcgc tgcgtcgcgg gctttggagc ggcgcagggc aacgagccga    20280

tcgctgatcg tggaaacgat aggcctatgc catgcgggtc aaggcgactt ccggcaagct    20340

atacgcgccc taggagtgcg gttggaacgt tggcccagcc agatactccc gatcacgagc    20400

aggacgccga tgatttgaag cgcactcagc gtctgatcca agaacaacca tcctagcaac    20460

acggcggtcc ccgggctgag aaagcccagt aaggaaacaa ctgtaggttc gagtcgcgag    20520

atcccccgga accaaaggaa gtaggttaaa cccgctccga tcaggccgag ccacgccagg    20580

ccgagaacat tggttcctgt aggcatcggg attggcggat caaacactaa agctactgga    20640

acgagcagaa gtcctccggc cgccagttgc caggcggtaa aggtgagcag aggcacggga    20700

ggttgccact tgcgggtcag cacggttccg aacgccatgg aaaccgcccc cgccaggccc    20760

gctgcgacgc cgacaggatc tagcgctgcg tttggtgtca acaccaacag cgccacgccc    20820

gcagttccgc aaatagcccc caggaccgcc atcaatcgta tcgggctacc tagcagagcg    20880

gcagagatga acacgaccat cagcggctgc acagcgccta ccgtcgccgc gaccccgccc    20940

ggcaggcggt agaccgaaat aaacaacaag ctccagaata gcgaaatatt aagtgcgccg    21000

aggatgaaga tgcgcatcca ccagattccc gttggaatct gtcggacgat catcacgagc    21060

aataaacccg ccggcaacgc ccgcagcagc ataccggcga cccctcggcc tcgctgttcg    21120

ggctccacga aaacgccgga cagatgcgcc ttgtgagcgt ccttgggccc gtcctcctgt    21180

ttgaagaccg acagcccaat gatctcgccg tcgatgtagg cgccgaatgc cacggcatct    21240

cgcaaccgtt cagcgaacgc ctccatgggc tttttctcct cgtgctcgta aacggacccg    21300

aacatctctg gagctttctt cagggccgac aatcggatct cgcggaaatc ctgcacgtcg    21360

gccgctccaa gccgtcgaat ctgagcctta atcacaattg tcaattttaa tcctctgttt    21420

atcggcagtt cgtagagcgc gccgtgcgtc ccgagcgata ctgagcgaag caagtgcgtc    21480

gagcagtgcc cgcttgttcc tgaaatgcca gtaaagcgct ggctgctgaa cccccagccg    21540

gaactgaccc cacaaggccc tagcgtttgc aatgcaccag gtcatcattg acccaggcgt    21600

gttccaccag gccgctgcct cgcaactctt cgcaggcttc gccgacctgc tcgcgccact    21660

tcttcacgcg ggtggaatcc gatccgcaca tgaggcggaa ggtttccagc ttgagcgggt    21720

acggctcccg gtgcgagctg aaatagtcga acatccgtcg ggccgtcggc gacagcttgc    21780

ggtacttctc ccatatgaat ttcgtgtagt ggtcgccagc aaacagcacg acgatttcct    21840

cgtcgatcag gacctggcaa cgggacgttt tcttgccacg gtccaggacg cggaagcggt    21900

gcagcagcga caccgattcc aggtgcccaa cgcggtcgga cgtgaagccc atcgccgtcg    21960

cctgtaggcg cgacaggcat tcctcggcct tcgtgtaata ccggccattg atcgaccagc    22020

ccaggtcctg gcaaagctcg tagaacgtga aggtgatcgg ctcgccgata ggggtgcgct    22080

tcgcgtactc caacacctgc tgccacacca gttcgtcatc gtcggcccgc agctcgacgc    22140

cggtgtaggt gatcttcacg tccttgttga cgtggaaaat gaccttgttt tgcagcgcct    22200

cgcgcgggat tttcttgttg cgcgtggtga acagggcaga gcgggccgtg tcgtttggca    22260

tcgctcgcat cgtgtccggc cacggcgcaa tatcgaacaa ggaaagctgc atttccttga    22320

tctgctgctt cgtgtgtttc agcaacgcgg cctgcttggc ctcgctgacc tgttttgcca    22380
```

-continued

```
ggtcctcgcc ggcggttttt cgcttcttgg tcgtcatagt tcctcgcgtg tcgatggtca   22440

tcgacttcgc caaacctgcc gcctcctgtt cgagacgacg cgaacgctcc acggcggccg   22500

atggcgcggg cagggcaggg ggagccagtt gcacgctgtc gcgctcgatc ttggccgtag   22560

cttgctggac catcgagccg acggactgga aggtttcgcg gggcgcacgc atgacggtgc   22620

ggcttgcgat ggtttcggca tcctcggcgg aaaaccccgc gtcgatcagt tcttgcctgt   22680

atgccttccg gtcaaacgtc cgattcattc accctccttg cgggattgcc ccgactcacg   22740

ccggggcaat gtgcccttat tcctgatttg acccgcctgg tgccttggtg tccagataat   22800

ccaccttatc ggcaatgaag tcggtcccgt agaccgtctg gccgtccttc cgtacttgg    22860

tattccgaat cttgccctgc acgaatacca gcgacccctt gcccaaatac ttgccgtggg   22920

cctcggcctg agagccaaaa cacttgatgc ggaagaagtc ggtgcgctcc tgcttgtcgc   22980

cggcatcgtt gcgccactct tcattaaccg ctatatcgaa aattgcttgc ggcttgttag   23040

aattgccatg acgtacctcg gtgtcacggg taagattacc gataaactgg aactgattat   23100

ggctcatatc gaaagtctcc ttgagaaagg agactctagt ttagctaaac attggttccg   23160

ctgtcaagaa cttttagcggc taaaattttg cgggccgcga ccaaaggtgc gaggggcggc   23220

ttccgctgtg tacaaccaga tatttttcac caacatcctt cgtctgctcg atgagcgggg   23280

catgacgaaa catgagctgt cggagagggc aggggtttca atttcgtttt tatcagactt   23340

aaccaacggt aaggccaacc cctcgttgaa ggtgatggag gccattgccg acgccctgga   23400

aactcccta cctcttctcc tggagtccac cgaccttgac cgcgaggcac tcgcggagat   23460

tgcgggtcat cctttcaaga gcagcgtgcc gcccggatac gaacgcatca gtgtggtttt   23520

gccgtcacat aaggcgttta tcgtaaagaa atggggcgac gacacccgaa aaaagctgcg   23580

tggaaggctc tgacgccaag ggttagggct tgcacttcct tctttagccg ctaaaacggc   23640

cccttctctg cgggccgtcg gctcgcgcat catatcgaca tcctcaacgg aagccgtgcc   23700

gcgaatggca tcgggcgggt gcgctttgac agttgttttc tatcagaacc cctacgtcgt   23760

gcggttcgat tagctgtttg tcttgcaggc taaacacttt cggtatatcg tttgcctgtg   23820

cgataatgtt gctaatgatt tgttgcgtag gggttactga aaagtgagcg ggaaagaaga   23880

gtttcagacc atcaaggagc gggccaagcg caagctggaa cgcgacatgg gtgcggacct   23940

gttggccgcg ctcaacgacc cgaaaaccgt tgaagtcatg ctcaacgcgg acggcaaggt   24000

gtggcacgaa cgccttggcg agccgatgcg gtacatctgc gacatgcggc ccagccagtc   24060

gcaggcgatt atagaaacgg tggccggatt ccacggcaaa gaggtcacgc ggcattcgcc   24120

catcctggaa ggcgagttcc ccttggatgg cagccgcttt gccggccaat gccgccggt    24180

cgtggccgcg ccaacctttg cgatccgcaa gcgcgcggtc gccatcttca cgctggaaca   24240

gtacgtcgag gcgggcatca tgacccgcga gcaatacgag gtcattaaaa gcgccgtcgc   24300

ggcgcatcga aacatcctcg tcattggcgg tactggctcg gcaagaccacgctcgtcaa     24360

cgcgatcatc aatgaaatgg tcgccttcaa cccgtctgag cgcgtcgtca tcatcgagga   24420

caccggcgaa atccagtgcg ccgcagagaa cgccgtccaa taccacacca gcatcgacgt   24480

ctcgatgacg ctgctgctca agacaacgct gcgtatgcgc cccgaccgca tcctggtcgg   24540

tgaggtacgt ggccccgaag cccttgatct gttgatggcc tggaacaccg ggcatgaagg   24600

aggtgccgcc accctgcacg caaacaaccc caaagcgggc ctgagccggc tcgccatgct   24660
```

-continued

```
tatcagcatg caccoggatt caccgaaacc cattgagccg ctgattggcg aggcggttca   24720

tgtggtcgtc catatcgcca ggacccctag cggccgtcga gtgcaagaaa ttctcgaagt   24780

tcttggttac gagaacggcc agtacatcac caaaaccctg taaggagtat ttccaatgac   24840

aacggctgtt ccgttccgtc tgaccatgaa tcgcggcatt ttgttctacc ttgccgtgtt   24900

cttcgttctc gctctcgcgt tatccgcgca tccggcgatg gcctcggaag gcaccggcgg   24960

cagcttgcca tatgagagct ggctgacgaa cctgcgcaac tccgtaaccg ccccggtggc   25020

cttcgcgctg tccatcatcg gcatcgtcgt cgccggcggc gtgctgatct tcggcggcga   25080

actcaacgcc ttcttccgaa ccctgatctt cctggttctg gtgatggcgc tgctggtcgg   25140

cgcgcagaac gtgatgagca ccttcttcgg tcgtggtgcc gaaatcgcgg ccctcggcaa   25200

cggggcgctg caccaggtgc aagtcgcggc ggcggatgcc gtgcgtgcgg tagcggctgg   25260

acggctcgcc taatcatggc tctgcgcacg atccccatcc gtcgcgcagg caaccgagaa   25320

aacctgttca tgggtggtga tcgtgaactg gtgatgttct cgggcctgat ggcgtttgcg   25380

ctgattttca gcgcccaaga gctgcgggcc accgtggtcg gtctgatcct gtggttcggg   25440

gcgctctatg cgttccgaat catggcgaag gccgatccga agatgcggtt cgtgtacctg   25500

cgtcaccgcc ggtacaagcc gtattacccg gcccgctcga ccccgttccg cgagaacacc   25560

aatagccaag ggaagcaata ccgatgatcc aagcaattgc gattgcaatc gcgggcctcg   25620

gcgcgcttct gttgttcatc ctctttgccc gcatccgcgc ggtcgatgcc gaactgaaac   25680

tgaaaaagca tcgttccaag gacgccggcc tggccgatct gctcaactac gccgctgtcg   25740

tcgatgacgg cgtaatcgtg ggcaagaacg gcagctttat ggctgcctgg ctgtacaagg   25800

gcgatgacaa cgcaagcagc accgaccagc agcgcgaagt agtgtccgcc cgcatcaacc   25860

aggccctcgc gggcctggga agtgggtgga tgatccatgt ggacgccgtg cggcgtcctg   25920

ctccgaacta cgcggagcgg ggcctgtcgg cgttccctga ccgtctgacg gcagcgattg   25980

aagaagagcg ctcggtcttg ccttgctcgt cggtgatgta cttcaccagc tccgcgaagt   26040

cgctcttctt gatggagcgc atggggacgt gcttggcaat cacgcgcacc ccccggccgt   26100

tttagcggct aaaaaagtca tggctctgcc ctcgggcgga ccacgcccat catgaccttg   26160

ccaagctcgt cctgcttctc ttcgatcttc gccagcaggg cgaggatcgt ggcatcaccg   26220

aaccgcgccg tgcgcgggtc gtcggtgagc cagagtttca gcaggccgcc caggcggccc   26280

aggtcgccat tgatgcgggc cagctcgcgg acgtgctcat agtccacgac gcccgtgatt   26340

ttgtagccct ggccgacggc cagcaggtag gccgacaggc tcatgccggc cgccgccgcc   26400

tttttcctcaa tcgctcttcg ttcgtctgga aggcagtaca ccttgatagg tgggctgccc   26460

ttcctggttg gcttggtttc atcagccatc cgcttgccct catctgttac gccggcggta   26520

gccggccagc ctcgcagagc aggattcccg ttgagcaccg ccaggtgcga ataagggaca   26580

gtgaagaagg aacacccgct cgcgggtggg cctacttcac ctatcctgcc cggctgacgg   26640

cgttggatac accaaggaaa gtctacacga acccttaggc aaaatcctgt atatcgtgcg   26700

aaaaaggatg gatataccga aaaaatcgct ataatgaccc cgaagcaggg ttatgcagcg   26760

gaaaagcgct gcttccctgc tgttttgtgg aatatctacc gactggaaac aggcaaatgc   26820

aggaaattac tgaactgagg ggacaggcga gagacgatgc caaagagcta caccgacgag   26880

ctggccgagt gggttgaatc ccgcgcggcc aagaagcgcc ggcgtgatga ggctgcggtt   26940
```

-continued

```
gcgttcctgg cggtgagggc ggatgtcgag gcggcgttag cgtccggcta tgcgctcgtc   27000

accatttggg agcacatgcg ggaaacgggg aaggtcaagt tctcctacga gacgttccgc   27060

tcgcacgcca ggcggcacat caaggccaag cccgccgatg tgcccgcacc gcaggccaag   27120

gctgcggaac ccgcgccggc acccaagacg ccggagccac ggcggccgaa gcagggggc    27180

aaggctgaaa agccggcccc cgctgcggcc ccgaccggct tcaccttcaa cccaacaccg   27240

gacaaaaagg atctactgta atggcgaaaa ttcacatggt tttgcagggc aagggcgggg   27300

tcggcaagtc ggccatcgcc gcgatcattg cgcagtacaa gatggacaag gggcagacac   27360

ccttgtgcat cgacaccgac ccggtgaacg cgacgttcga gggctacaag gccctgaacg   27420

tccgccggct gaacatcatg gccggcgacg aaattaactc gcgcaacttc gacaccctgg   27480

tcgagctgat tgcgccgacc aaggatgacg tggtgatcga caacggtgcc agctcgttcg   27540

tgcctctgtc gcattacctc atcagcaacc aggtgccggc tctgctgcaa gaaatggggc   27600

atgagctggt catccatacc gtcgtcaccg gcggccaggc tctcctggac acggtgagcg   27660

gcttcgccca gctcgccagc cagttcccgg ccgaagcgct tttcgtggtc tggctgaacc   27720

cgtattgggg gcctatcgag catgagggca agagctttga gcagatgaag gcgtacacgg   27780

ccaacaaggc ccgcgtgtcg tccatcatcc agattccggc cctcaaggaa gaaacctacg   27840

gccgcgattt cagcgacatg ctgcaagagc ggctgacgtt cgaccaggcg ctggccgatg   27900

aatcgctcac gatcatgacg cggcaacgcc tcaagatcgt gcggcgcggc ctgtttgaac   27960

agctcgacgc ggcggccgtg ctatgagcga ccagattgaa gagctgatcc gggagattgc   28020

ggccaagcac ggcatcgccg tcggccgcga cgacccggtg ctgatcctgc ataccatcaa   28080

cgcccggctc atggccgaca gtgcggccaa gcaagaggaa atccttgccg cgttcaagga   28140

agagctggaa gggatcgccc atcgttgggg cgaggacgcc aaggccaaag cggagcggat   28200

gctgaacgcg gccctggcgg ccagcaagga cgcaatggcg aaggtaatga aggacagcgc   28260

cgcgcaggcg gccgaagcga tccgcaggga aatcgacgac ggccttggcc gccagctcgc   28320

ggccaaggtc gcggacgcgc ggcgcgtggc gatgatgaac atgatcgccg gcggcatggt   28380

gttgttcgcg gccgccctgg tggtgtgggc ctcgttatga atcgcagagg cgcagatgaa   28440

aaagcccggc gttgccgggc tttgtttttg cgttagctgg gcttgtttga caggcccaag   28500

ctctgactgc gcccgcgctc gcgctcctgg gcctgtttct tctcctgctc ctgcttgcgc   28560

atcagggcct ggtgccgtcg ggctgcttca cgcatcgaat cccagtcgcc ggccagctcg   28620

ggatgctccg cgcgcatctt gcgcgtcgcc agttcctcga tcttgggcgc gtgaatgccc   28680

atgccttcct tgatttcgcg caccatgtcc agccgcgtgt gcagggtctg caagcgggct   28740

tgctgttggg cctgctgctg ctgccaggcg gcctttgtac gcggcaggga cagcaagccg   28800

ggggcattgg actgtagctg ctgcaaacgc gcctgctgac ggtctacgag ctgttctagg   28860

cggtcctcga tgcgctccac ctggtcatgc tttgcctgca cgtagagcgc aagggtctgc   28920

tggtaggtct gctcgatggg cgcggattct aagagggcct gctgttccgt ctcggcctcc   28980

tgggccgcct gtagcaaatc ctcgccgctg ttgccgctgg actgctttac tgccggggac   29040

tgctgttgcc ctgctcgcgc cgtcgtcgca gttcggcttg ccccactcg attgactgct     29100

tcatttcgag ccgcagcgat gcgatctcgg attgcgtcaa cggacggggc agcgcggagg   29160

tgtccggctt ctccttgggt gagtcggtcg atgccatagc caaaggtttc cttccaaaat   29220
```

-continued

```
gcgtccattg ctggaccgtg tttctcattg atgcccgcaa gcatcttcgg cttgaccgcc    29280

aggtcaagcg cgccttcatg ggcggtcatg acggacgccg ccatgacctt gccgccgttg    29340

ttctcgatgt agccgcgtaa tgaggcaatg gtgccgccca tcgtcagcgt gtcatcgaca    29400

acgatgtact tctggccggg gatcacctcc ccctcgaaag tcgggttgaa cgccaggcga    29460

tgatctgaac cggctccggt tcgggcgacc ttctcccgct gcacaatgtc cgtttcgacc    29520

tcaaggccaa ggcggtcggc cagaacgacc gccatcatgg ccggaatctt gttgttcccc    29580

gccgcctcga cggcgaggac tggaacgatg cggggcttgt cgtcgccgat cagcgtcttg    29640

agctgggcaa cagtgtcgtc cgaaatcagg cgctcgacca aattaagcgc cgcttccgcg    29700

tcgccctgct tcgcagcctg gtattcaggc tcgttggtca agaaccaag gtcgccgttg     29760

cgaaccacct tcgggaagtc tccccacggt gcgcgctcgg ctctgctgta gctgctcaag    29820

acgcctccct ttttagccgc taaaactcta acgagtgcgc ccgcgactca acttgacgct    29880

ttcggcactt acctgtgcct tgccacttgc gtcataggtg atgctttcg cactcccgat     29940

ttcaggtact ttatcgaaat ctgaccgggc gtgcattaca aagttcttcc ccacctgttg    30000

gtaaatgctg ccgctatctg cgtggacgat gctgccgtcg tggcgctgcg acttatcggc    30060

cttttgggcc atatagatgt tgtaaatgcc aggtttcagg gccccggctt tatctacctt    30120

ctggttcgtc catgcgcctt ggttctcggt ctggacaatt ctttgcccat tcatgaccag    30180

gaggcggtgt ttcattgggt gactcctgac ggttgcctct ggtgttaaac gtgtcctggt    30240

cgcttgccgg ctaaaaaaaa gccgacctcg gcagttcgag gccggctttc cctagagccg    30300

ggcgcgtcaa ggttgttcca tctattttag tgaactgcgt tcgatttatc agttactttc    30360

ctcccgcttt gtgtttcctc ccactcgttt ccgcgtctag ccgacccctc aacatagcgg    30420

cctcttcttg ggctgccttt gcctcttgcc gcgcttcgtc acgctcggct tgcaccgtcg    30480

taaagcgctc ggcctgcctg gccgcctctt gcgccgccaa cttcctttgc tcctggtggg    30540

cctcggcgtc ggcctgcgcc ttcgctttca ccgctgccaa ctccgtgcgc aaactctccg    30600

cttcgcgcct ggtggcgtcg cgctcgccgc gaagcgcctg catttcctgg ttggccgcgt    30660

ccagggtctt gcggctctct tctttgaatg cgcgggcgtc ctggtgagcg tagtccagct    30720

cggcgcgcag ctcctgcgct cgacgctcca cctcgtcggc ccgctgcgtc gccagcgcgg    30780

cccgctgctc ggctcctgcc agggcggtgc gtgcttcggc cagggcttgc cgctggcgtg    30840

cggccagctc ggccgcctcg gcggcctgct gctctagcaa tgtaacgcgc gcctgggctt    30900

cttccagctc gcgggcctgc gcctcgaagg cgtcggccag ctcccgcgc acggcttcca     30960

actcgttgcg ctcacgatcc cagccggctt gcgctgcctg caacgattca ttggcaaggg    31020

cctgggcggc ttgccagagg gcggccacgg cctggttgcc ggcctgctgc accgcgtccg    31080

gcacctggac tgccagcggg gcggcctgcg ccgtgcgctg gcgtcgccat cgcgcatgc     31140

cggcgctggc gtcgttcatg ttgacgcggg cggccttacg cactcatcc acggtcggga     31200

agttctcccg gtcgccttgc tcgaacagct cgtccgcagc cgcaaaaatg cggtcgcgcg    31260

tctctttgtt cagttccatg ttggctccgg taattggtaa gaataataat actcttacct    31320

accttatcag cgcaagagtt tagctgaaca gttctcgact taacggcagg ttttttagcg    31380

gctgaagggc aggcaaaaaa agccccgcac ggtcggcggg ggcaaagggt cagcgggaag    31440

gggattagcg ggcgtcgggc ttcttcatgc gtcggggccg cgcttcttgg gatggagcac    31500
```

-continued

```
gacgaagcgc gcacgcgcat cgtcctcggc cctatcggcc cgcgtcgcgg tcaggaactt   31560

gtcgcgcgct aggtcctccc tggtgggcac caggggcatg aactcggcct gctcgatgta   31620

ggtccactcc atgaccgcat cgcagtcgag gccgcgttcc ttcaccgtct cttgcaggtc   31680

gcggtacgcc cgctcgttga gcggctggta acgggccaat tggtcgtaaa tggctgtcgg   31740

ccatgagcgg cctttcctgt tgagccagca gccgacgacg aagccggcaa tgcaggcccc   31800

tggcacaacc aggccgacgc cggggcagg ggatggcagc agctcgccaa ccaggaaccc   31860

cgccgcgatg atgccgatgc cggtcaacca gcccttgaaa ctatccggcc ccgaaacacc   31920

cctgcgcatt gcctggatgc tgcgccggat agcttgcaac atcaggagcc gtttctttg   31980

ttcgtcagtc atggtccgcc ctcaccagtt gttcgtatcg gtgtcggacg aactgaaatc   32040

gcaagagctg ccggtatcgg tccagccgct gtccgtgtcg ctgctgccga agcacggcga   32100

ggggtccgcg aacgccgcag acggcgtatc cggccgcagc gcatcgccca gcatggcccc   32160

ggtcagcgag ccgccggcca ggtagcccag catggtgctg ttggtcgccc cggccaccag   32220

ggccgacgtg acgaaatcgc cgtcattccc tctggattgt tcgctgctcg gcggggcagt   32280

gcgccgcgcc ggcggcgtcg tggatggctc gggttggctg gcctgcgacg gccggcgaaa   32340

ggtgcgcagc agctcgttat cgaccggctg cggcgtcggg gccgccgcct tgcgctgcgg   32400

tcggtgttcc ttcttcggct cgcgcagctt gaacagcatg atcgcggaaa ccagcagcaa   32460

cgccgcgcct acgcctcccg cgatgtagaa cagcatcgga ttcattcttc ggtcctcctt   32520

gtagcggaac cgttgtctgt gcggcgcggg tggcccgcgc cgctgtcttt ggggatcagc   32580

cctcgatgag cgcgaccagt ttcacgtcgg caaggttcgc ctcgaactcc tggccgtcgt   32640

cctcgtactt caaccaggca tagccttccg ccggcggccg acggttgagg ataaggcggg   32700

cagggcgctc gtcgtgctcg acctggacga tggccttttt cagcttgtcc gggtccggct   32760

ccttcgcgcc cttttccttg gcgtccttac cgtcctggtc gccgtcctcg ccgtcctggc   32820

cgtcgccggc ctccgcgtca cgctcggcat cagtctggcc gttgaaggca tcgacggtgt   32880

tgggatcgcg gcccttctcg tccaggaact cgcgcagcag cttgaccgtg ccgcgcgtga   32940

tttcctgggt gtcgtcgtca agccacgcct cgacttcctc cgggcgcttc ttgaaggccg   33000

tcaccagctc gttcaccacg gtcacgtcgc gcacgcggcc ggtgttgaac gcatcggcga   33060

tcttctccgg caggtccagc agcgtgacgt gctgggtgat gaacgccggc gacttgccga   33120

tttccttggc gatatcgcct ttcttcttgc ccttcgccag ctcgcggcca atgaagtcgg   33180

caatttcgcg cggggtcagc tcgttgcgtt gcaggttctc gataacctgg tcggcttcgt   33240

tgtagtcgtt gtcgatgaac gccgggatgg acttcttgcc ggcccacttc gagccacggt   33300

agcggcgggc gccgtgattg atgatatagc ggcccggctg ctcctggttc tcgcgcaccg   33360

aaatgggtga cttcaccccg cgctctttga tcgtggcacc gatttccgcg atgctctccg   33420

gggaaaagcc ggggttgtcg gccgtccgcg gctgatgcgg atcttcgtcg atcaggtcca   33480

ggtccagctc gatagggccg gaaccgccct gagacgccgc aggagcgtcc aggaggctcg   33540

acaggtcgcc gatgctatcc aaccccaggc cggacggctg cgccgcgcct gcggcttcct   33600

gagcggccgc agcggtgttt ttcttggtgg tcttggcttg agccgcagtc attgggaaat   33660

ctccatcttc gtgaacacgt aatcagccag ggcgcgaacc tctttcgatg ccttgcgcgc   33720

ggccgttttc ttgatcttcc agaccggcac accggatgcg agggcatcgg cgatgctgct   33780
```

-continued

```
gcgcaggcca acggtggccg gaatcatcat cttggggtac gcggccagca gctcggcttg   33840

gtggcgcgcg tggcgcggat tccgcgcatc gaccttgctg ggcaccatgc caaggaattg   33900

cagcttggcg ttcttctggc gcacgttcgc aatggtcgtg accatcttct tgatgccctg   33960

gatgctgtac gcctcaagct cgatggggga cagcacatag tcggccgcga agagggcggc   34020

cgccaggccg acgccaaggg tcggggccgt gtcgatcagg cacacgtcga agccttggtt   34080

cgccagggcc ttgatgttcg ccccgaacag ctcgcgggcg tcgtccagcg acagccgttc   34140

ggcgttcgcc agtaccgggt tggactcgat gagggcgagg cgcgcggcct ggccgtcgcc   34200

ggctgcgggt gcggtttcgg tccagccgcc ggcaggaca gcgccgaaca gcttgcttgc   34260

atgcaggccg gtagcaaagt ccttgagcgt gtaggacgca ttgccctggg ggtccaggtc   34320

gatcacggca acccgcaagc cgcgctcgaa aaagtcgaag gcaagatgca caagggtcga   34380

agtcttgccg acgccgcctt tctggttggc cgtgaccaaa gttttcatcg tttggtttcc   34440

tgtttttttct tggcgtccgc ttcccacttc cggacgatgt acgcctgatg ttccggcaga   34500

accgccgtta cccgcgcgta cccctcgggc aagttcttgt cctcgaacgc ggcccacacg   34560

cgatgcaccg cttgcgacac tgcgcccctg gtcagtccca gcgacgttgc gaacgtcgcc   34620

tgtggcttcc catcgactaa gacgcccgc gctatctcga tggtctgctg ccccacttcc   34680

agccctggga tcgcctcctg gaactggctt tcggtaagcc gtttcttcat ggataacacc   34740

cataatttgc tccgcgcctt ggttgaacat agcggtgaca gccgccagca catgagagaa   34800

gtttagctaa acatttctcg cacgtcaaca cctttagccg ctaaaactcg tccttggcgt   34860

aacaaaacaa aagcccggaa accgggcttt cgtctcttgc cgcttatggc tctgcacccg   34920

gctccatcac caacaggtcg cgcacgcgct tcactcggtt gcggatcgac actgccagcc   34980

caacaaagcc ggttgccgcc gccgccagga tcgcgccgat gatgccggcc acaccggcca   35040

tcgcccacca ggtcgccgcc ttccggttcc attcctgctg gtactgcttc gcaatgctgg   35100

acctcggctc accataggct gaccgctcga tggcgtatgc cgcttctccc cttggcgtaa   35160

aacccagcgc cgcaggcggc attgccatgc tgcccgccgc tttcccgacc acgacgcgcg   35220

caccaggctt gcggtccaga ccttcggcca cggcgagctg cgcaaggaca taatcagccg   35280

ccgacttggc tccacgcgcc tcgatcagct cttgcactcg cgcgaaatcc ttggcctcca   35340

cggccgccat gaatcgcgca cgcggcgaag gctccgcagg gccggcgtcg tgatcgccgc   35400

cgagaatgcc cttcaccaag ttcgacgaca cgaaaatcat gctgacggct atcaccatca   35460

tgcagacgga tcgcacgaac ccgctgaatt gaacacgagc acggcacccg cgaccactat   35520

gccaagaatg cccaaggtaa aaattgccgg ccccgccatg aagtccgtga atgccccgac   35580

ggccgaagtg aagggcaggc cgccacccag gccgccgccc tcactgcccg gcacctggtc   35640

gctgaatgtc gatgccagca cctgcggcac gtcaatgctt ccgggcgtcg cgctcgggct   35700

gatcgcccat cccgttactg ccccgatccc ggcaatggca aggactgcca gcgctgccat   35760

ttttggggtg aggccgttcg cggccgaggg gcgcagcccc tgggggatg ggaggcccgc   35820

gttagcgggc cgggagggtt cgagaagggg gggcacccccc cttcggcgtg cgcggtcacg   35880

cgcacagggc gcagccctgg ttaaaaacaa ggtttataaa tattggttta aaagcaggtt   35940

aaaagacagg ttagcggtgg ccgaaaaacg ggcggaaacc cttgcaaatg ctggattttc   36000

tgcctgtgga cagcccctca aatgtcaata ggtgcgcccc tcatctgtca gcactctgcc   36060
```

-continued

```
cctcaagtgt caaggatcgc gcccctcatc tgtcagtagt cgcgcccctc aagtgtcaat   36120

accgcagggc acttatcccc aggcttgtcc acatcatctg tgggaaactc gcgtaaaatc   36180

aggcgttttc gccgatttgc gaggctggca agctccacgt cgccggccga aatcgagcct   36240

gcccctcatc tgtcaacgcc gcgccgggtg agtcggcccc tcaagtgtca acgtccgccc   36300

ctcatctgtc agtgagggcc aagtttttccg cgaggtatcc acaacgccgg cggccgcggt   36360

gtctcgcaca cggcttcgac ggcgtttctg gcgcgtttgc agggccatag acggccgcca   36420

gcccagcggc gagggcaacc agcccggtga gcgtcggaaa ggcgctggaa gccccgtagc   36480

gacgcggaga gggcgagac aagccaaggg cgcaggctcg atgcgcagca cgacatagcc   36540

ggttctcgca aggacgagaa tttccctgcg gtgcccctca agtgtcaatg aaagtttcca   36600

acgcgagcca ttcgcgagag ccttgagtcc acgctagatg agagctttgt tgtaggtgga   36660

ccagttggtg attttgaact tttgctttgc cacggaacgg tctgcgttgt cgggaagatg   36720

cgtgatctga tccttcaact cagcaaaagt tcgatttatt caacaaagcc acgttgtgtc   36780

tcaaaatctc tgatgttaca ttgcacaaga taaaaatata tcatcatgaa caataaaact   36840

gtctgcttac ataaacagta atacaagggg tgttatgagc catattcaac gggaaacgtc   36900

ttgctcgac                                                           36909
```

<210> SEQ ID NO 33
<211> LENGTH: 825
<212> TYPE: DNA
<213> ORGANISM: Zea Mays

<400> SEQUENCE: 33

```
aaatccttac agaattgctg tagtttcata gtgctagatg tggacagcaa agcgccgctg     60

tatgcttctg ctttttctttt ttggtgtgtg tagccacatc ctttgttcct gcccggcgcc    120

atcccacttg gttgtttttt tttatgattg aaagccttca tgcttcctcg gtcaatcacc    180

ggtgcgcact gggagcatcg ccggaaaaaa aattcttcgg ctaagagtaa cttctttctc    240

cttttcttct ctgatctcgc gagcagtgct gataacgtgt tgtaatctac ttagcggtaa    300

cgagattgag agagacaaaa tgacagaact attgtcttta ttgcagagtg tcatgtattt    360

atacagggga tacaaagtct cccaaggggt gtgtcccttg ggagtaactg ccagttgatc    420

acaggacaat attttgtaac aaaacgtaca catcgtcaaa atagcgaggc atgaaactgg    480

ccttggccat ggacgcgtga agcgcgccat gcgttggata tgtggtcaat aagtatatac    540

aatacaatgt ttaacagagc tgatagtact gctttggcac attttttgtcc acgcttcatg    600

agagataaaa cacctgcacg taaattcaca tgctgcactg aaggcccgat cactgaggag    660

cgaactgccg taactccctt ctatatatac ccccagtccc tgtttcagtt ttcgtcaagc    720

tagcagcacc aagttgtcga tcacttgcct gctcttgagc tcgattaagc tatcatcagc    780

tacagcatcc gatcccaaac tgcaactgta gcagcgacaa ctgcc                     825
```

<210> SEQ ID NO 34
<211> LENGTH: 860
<212> TYPE: DNA
<213> ORGANISM: Zea mays

<400> SEQUENCE: 34

```
ctggtaatta ttggctgtag gattctaaac agagcctaaa tagctggaat agctctagcc     60
```

-continued

```
ctcaatccaa actaatgata tctatactta tgcaactcta aatttttatt ctaaaagtaa    120

tatttcattt ttgtcaacga gattctctac tctattccac aatcttttga agcaatattt    180

accttaaatc tgtactctat accaataatc atatattcta ttatttattt ttatctctct    240

cctaaggagc atcccctat gtctgcatgg ccccgcctc gggtcccaat ctcttgctct    300

gctagtagca cagaagaaaa cactagaaat gacttgcttg acttagagta tcagataaac    360

atcatgttta cttaactta atttgtatcg gtttctacta tttttataat attttttgtct    420

ctatagatac tacgtgcaac agtataatca acctagttta atccagagcg aaggattttt    480

tactaagtac gtgactccat atgcacagcg ttccttttat ggttcctcac tgggcacagc    540

ataaacgaac cctgtccaat gttttcagcg cgaacaaaca gaaattccat cagcgaacaa    600

acaacataca tgcgagatga aaataaataa taaaaaaagc tccgtctcga taggccggca    660

cgaatcgaga gcctccatag ccagtttttt ccatcggaac ggcggttcgc gcacctaatt    720

atatgcacca cacgcctata aagccaacca acccgtcgga ggggcgcaag ccagacagaa    780

gacagcccgt cagcccctct cgttttttcat ccgccttcgc ctccaaccgc gtgcgctcca    840

cgcctcctcc aggaaagcga                                                860
```

```
<210> SEQ ID NO 35
<211> LENGTH: 896
<212> TYPE: DNA
<213> ORGANISM: Zea mays

<400> SEQUENCE: 35

gtgcagcgtg acccggtcgt gcccctctct agagataatg agcattgcat gtctaagtta     60

taaaaaatta ccacatattt tttttgtcac acttgtttga agtgcagttt atctatcttt    120

atacatatat ttaaacttta ctctacgaat aatataatct atagtactac aataatatca    180

gtgtttttaga gaatcatata aatgaacagt tagacatggt ctaaaggaca attgagtatt    240

ttgacaacag gactctacag ttttatcttt ttagtgtgca tgtgttctcc ttttttttttg    300

caaatagctt cacctatata atacttcatc cattttatta gtacatccat ttagggttta    360

gggttaatgg tttttataga ctaatttttt tagtacatct attttattct attttagcct    420

ctaaattaag aaaactaaaa ctctatttta gtttttttat ttaataattt agatataaaa    480

tagaataaaa taaagtgact aaaaattaaa caaatacccct ttaagaaatt aaaaaaacta    540

aggaaacatt tttcttgttt cgagtagata atgccagcct gttaaacgcc gtcgacgagt    600

ctaacggaca ccaaccagcg aaccagcagc gtcgcgtcgg gccaagcgaa gcagacggca    660

cggcatctct gtcgctgcct ctggaccccct ctcgagagtt ccgctccacc gttggacttg    720

ctccgctgtc ggcatccaga aattgcgtgg cggagcggca gacgtgagcc ggcacggcag    780

gcggcctcct cctcctctca cggcaccggc agctacgggg gattccttc ccaccgctcc    840

ttcgctttcc cttcctcgcc cgccgtaata aatagacacc ccctccacac cctctt        896
```

```
<210> SEQ ID NO 36
<211> LENGTH: 318
<212> TYPE: DNA
<213> ORGANISM: Zea mays

<400> SEQUENCE: 36

agacttgtcc atcttctgga ttggccaact taattaatgt atgaaataaa aggatgcaca     60
```

-continued

```
catagtgaca tgctaatcac tataatgtgg gcatcaaagt tgtgtgttat gtgtaattac    120

tagttatctg aataaaagag aaagagatca tccatatttc ttatcctaaa tgaatgtcac    180

gtgtctttat aattctttga tgaaccagat gcatttcatt aaccaaatcc atatacatat    240

aaatattaat catatataat taatatcaat tgggttagca aaacaaatct agtctaggtg    300

tgttttgcga attgcggc                                                  318


<210> SEQ ID NO 37
<211> LENGTH: 4678
<212> TYPE: DNA
<213> ORGANISM: artificial
<220> FEATURE:
<223> OTHER INFORMATION: vector

<400> SEQUENCE: 37

gaaaggccca gtcttccgac tgagcctttc gttttatttg atgcctggca gttccctact     60

ctcgcgttaa cgctagcatg gatgtttttc cagtcacgac gttgtaaaac gacggccagt    120

cttaagctcg ggcccgcgtt aacgctacca tggagctcca ataatgatt ttattttgac     180

tgatagtgac ctgttcgttg caacaaattg ataagcaatg ctttttttata atgccaactt    240

tgtatagaaa agttgggccg aattcgagct cggtacggcc agaatggccc ggaccgggtt    300

accgaattcg agctcggtac cctgggatcc ctggtaatta ttggctgtag gattctaaac    360

agagcctaaa tagctggaat agctctagcc ctcaatccaa actaatgata tctatactta    420

tgcaactcta aattttatt ctaaaagtaa tatttcattt ttgtcaacga gattctctac     480

tctattccac aatcttttga agcaatattt accttaaatc tgtactctat accaataatc    540

atatattcta ttatttattt ttatctctct cctaaggagc atcccctat gtctgcatgg     600

ccccgcctc gggtcccaat ctcttgctct gctagtagca cagaagaaaa cactagaaat     660

gacttgcttg acttagagta tcagataaac atcatgttta cttaacttta atttgtatcg     720

gtttctacta ttttttataat attttttgtct ctatagatac tacgtgcaac agtataatca    780

acctagttta atccagagcg aaggattttt tactaagtac gtgactccat atgcacagcg     840

ttccttttat ggttcctcac tgggcacagc ataaacgaac cctgtccaat gttttcagcg     900

cgaacaaaca gaaattccat cagcgaacaa acaacataca tgcgagatga aaataaataa     960

taaaaaaagc tccgtctcga taggccggca cgaatcgaga gcctccatag ccagtttttt    1020

ccatcggaac ggcggttcgc gcacctaatt atatgcacca cacgcctata aagccaacca    1080

acccgtcgga ggggcgcaag ccagacagaa gacagcccgt cagcccctct cgttttttcat   1140

ccgccttcgc ctccaaccgc gtgcgctcca cgcctcctcc aggaaagcga ggatctcccc    1200

caaatccacc cgtcggcacc tccgcttcaa ggtacgccgc tcgtcctccc cccccccccc    1260

tctctacctt ctctagatcg gcgttccggt ccatggttag ggcccggtag ttctacttct    1320

gttcatgttt gtgttagatc cgtgtttgtg ttagatccgt gctgctagcg ttcgtacacg    1380

gatgcgacct gtacgtcaga cacgttctga ttgctaactt gccagtgttt ctctttgggg    1440

aatcctggga tggctctagc cgttccgcag acgggatcga tttcatgatt ttttttgttt    1500

cgttgcatag ggtttggttt gccctttttcc tttatttcaa tatatgccgt gcacttgttt    1560

gtcgggtcat cttttcatgc tttttttttgt cttggttgtg atgatgtggt ctggttgggc    1620

ggtcgttcta gatcggagta gaattctgtt tcaaactacc tggtggattt attaattttg    1680
```

-continued

```
gatctgtatg tgtgtgccat acatattcat agttacgaat tgaagatgat ggatggaaat    1740

atcgatctag ataggtata catgttgatg cgggttttac tgatgcatat acagagatgc    1800

tttttgttcg cttggttgtg atgatgtggt gtggttgggc ggtcgttcat tcgttctaga    1860

tcggagtaga atactgtttc aaactacctg gtgtatttat taattttgga actgtatgtg    1920

tgtgtcatac atcttcatag ttacgagttt aagatggatg gaaatatcga tctaggatag    1980

gtatacatgt tgatgtgggt tttactgatg catatacatg atggcatatg cagcatctat    2040

tcatatgctc taaccttgag tacctatcta ttataataaa caagtatgtt ttataattat    2100

tttgatcttg atatacttgg atgatggcat atgcagcagc tatatgtgga ttttttttagc    2160

cctgccttca tacgctattt atttgcttgg tactgtttct tttgtcgatg ctcaccctgt    2220

tgtttggtgt tacttctgca ggtcgactct agaagcttgg tcacccggtc cgggcctaga    2280

aggccagctt caagtttgta caaaaaagtt gaacgagaaa cgtaaaatga tataaatatc    2340

aatatattaa attagatttt gcataaaaaa cagactacat aatactgtaa aacacaacat    2400

atgcagtcac tatgaatcaa ctacttagat ggtattagtg acctgtagaa ttcgagctct    2460

agagctgcag ggcggccgcg atatccccta tagtgagtcg tattacatgg tcatagctgt    2520

ttcctggcag ctctggcccg tgtctcaaaa tctctgatgt tacattgcac aagataaaaa    2580

tatatcatca tgaacaataa aactgtctgc ttacataaac agtaatacaa ggggtgttat    2640

gagccatatt caacgggaaa cgtcgaggcc gcgattaaat tccaacatgg atgctgattt    2700

atatgggtat aaatgggctc gcgataatgt cggcaatca ggtgcgacaa tctatcgctt    2760

gtatgggaag cccgatgcgc cagagttgtt tctgaaacat ggcaaaggta gcgttgccaa    2820

tgatgttaca gatgagatgg tcagactaaa ctggctgacg gaatttatgc ctcttccgac    2880

catcaagcat tttatccgta ctcctgatga tgcatggtta ctcaccactg cgatccccgg    2940

aaaaacagca ttccaggtat tagaagaata tcctgattca ggtgaaaata ttgttgatgc    3000

gctggcagtg ttcctgcgcc ggttgcattc gattcctgtt tgtaattgtc cttttaacag    3060

cgatcgcgta tttcgtctcg ctcaggcgca atcacgaatg aataacggtt tggttgatgc    3120

gagtgatttt gatgacgagc gtaatggctg gcctgttgaa caagtctgga agaaatgca    3180

taaacttttg ccattctcac cggattcagt cgtcactcat ggtgatttct cacttgataa    3240

ccttattttt gacgagggga aattaatagg ttgtattgat gttggacgag tcggaatcgc    3300

agaccgatac caggatcttg ccatcctatg gaactgcctc ggtgagtttt ctccttcatt    3360

acagaaacgg cttttttcaaa aatatggtat tgataatcct gatatgaata aattgcagtt    3420

tcatttgatg ctcgatgagt ttttctaatc agaattggtt aattggttgt aacactggca    3480

gagcattacg ctgacttgac gggacggcgc aagctcatga ccaaaatccc ttaacgtgag    3540

ttacgcgtcg ttccactgag cgtcagaccc cgtagaaaag atcaaaggat cttcttgaga    3600

tcctttttttt ctgcgcgtaa tctgctgctt gcaaacaaaa aaaccaccgc taccagcggt    3660

ggtttgtttg ccggatcaag agctaccaac tcttttttccg aaggtaactg gcttcagcag    3720

agcgcagata ccaaatactg tccttctagt gtagccgtag ttaggccacc acttcaagaa    3780

ctctgtagca ccgcctacat acctcgctct gctaatcctg ttaccagtgg ctgctgccag    3840

tggcgataag tcgtgtctta ccgggttgga ctcaagacga tagttaccgg ataaggcgca    3900

gcggtcgggc tgaacggggg gttcgtgcac acagcccagc ttggagcgaa cgacctacac    3960
```

-continued

```
cgaactgaga tacctacagc gtgagcattg agaaagcgcc acgcttcccg aagggagaaa    4020

ggcggacagg tatccggtaa gcggcagggt cggaacagga gagcgcacga gggagcttcc    4080

aggggggaaac gcctggtatc tttatagtcc tgtcgggttt cgccacctct gacttgagcg    4140

tcgattttttg tgatgctcgt caggggggcg gagcctatgg aaaaacgcca gcaacgcggc    4200

ctttttacgg ttcctggcct tttgctggcc ttttgctcac atgttctttc ctgcgttatc    4260

ccctgattct gtggataacc gtattaccgc ctttgagtga gctgataccg ctcgccgcag    4320

ccgaacgacc gagcgcagcg agtcagtgag cgaggaagcg gaagagcgcc caatacgcaa    4380

accgcctctc cccgcgcgtt ggccgattca ttaatgcagc tggcacgaca ggtttcccga    4440

ctggaaagcg ggcagtgagc gcaacgcaat taatacgcgt accgctagcc aggaagagtt    4500

tgtagaaacg caaaaaggcc atccgtcagg atggccttct gcttagtttg atgcctggca    4560

gtttatggcg ggcgtcctgc ccgccaccct ccgggccgtt gcttcacaac gttcaaatcc    4620

gctcccggcg gatttgtcct actcaggaga gcgttcaccg acaaacaaca gataaaac      4678
```

<210> SEQ ID NO 38
<211> LENGTH: 2991
<212> TYPE: DNA
<213> ORGANISM: artificial
<220> FEATURE:
<223> OTHER INFORMATION: vector

<400> SEQUENCE: 38

```
ctttcctgcg ttatcccctg attctgtgga taaccgtatt accgcctttg agtgagctga      60

taccgctcgc cgcagccgaa cgaccgagcg cagcgagtca gtgagcgagg aagcggaaga    120

gcgcccaata cgcaaaccgc ctctccccgc gcgttggccg attcattaat gcagctggca    180

cgacaggttt cccgactgga aagcgggcag tgagcgcaac gcaattaata cgcgtaccgc    240

tagccaggaa gagtttgtag aaacgcaaaa aggccatccg tcaggatggc cttctgctta    300

gtttgatgcc tggcagttta tggcgggcgt cctgcccgcc accctccggg ccgttgcttc    360

acaacgttca aatccgctcc cggcggattt gtcctactca ggagagcgtt caccgacaaa    420

caacagataa aacgaaaggc ccagtcttcc gactgagcct ttcgttttat ttgatgcctg    480

gcagttccct actctcgcgt taacgctagc atggatgttt cccagtcac gacgttgtaa    540

aacgacggcc agtcttaagc tcgggccctg cagctctaga gctcgaattc tacaggtcac    600

taataccatc taagtagttg gttcatagtg actgcatatg ttgtgtttta cagtattatg    660

tagtctgttt tttatgcaaa atctaattta atatattgat atttatatca ttttacgttt    720

ctcgttcaac tttcttgtac aaagtggccg ttaacggatc cagacttgtc catcttctgg    780

attggccaac ttaattaatg tatgaaataa aaggatgcac acatagtgac atgctaatca    840

ctataatgtg ggcatcaaag ttgtgtgtta tgtgtaatta ctagttatct gaataaaaga    900

gaaagagatc atccatattt cttatcctaa atgaatgtca cgtgtcttta taattctttg    960

atgaaccaga tgcatttcat taaccaaatc catatacata taaatattaa tcatatataa   1020

ttaatatcaa ttgggttagc aaaacaaatc tagtctaggt gtgttttgcg aattgcggca   1080

agcttgcggc cgccccgggc aactttatta tacaaagttg gcattataaa aaagcattgc   1140

ttatcaattt gttgcaacga acaggtcact atcagtcaaa ataaaatcat tatttggagc   1200

tccatggtag cgttaacgcg ccgcgatat cccctatagt gagtcgtatt acatggtcat   1260
```

-continued

```
agctgtttcc tggcagctct ggcccgtgtc tcaaaatctc tgatgttaca ttgcacaaga      1320

taaaaatata tcatcatgaa caataaaact gtctgcttac ataaacagta atacaagggg      1380

tgttatgagc catattcaac gggaaacgtc gaggccgcga ttaaattcca acatggatgc      1440

tgatttatat gggtataaat gggctcgcga taatgtcggg caatcaggtg cgacaatcta      1500

tcgcttgtat gggaagcccg atgcgccaga gttgtttctg aaacatggca aaggtagcgt      1560

tgccaatgat gttacagatg agatggtcag actaaactgg ctgacggaat ttatgcctct      1620

tccgaccatc aagcattttа tccgtactcc tgatgatgca tggttactca ccactgcgat      1680

cccccggaaaa acagcattcc aggtattaga agaatatcct gattcaggtg aaaatattgt      1740

tgatgcgctg gcagtgttcc tgcgccggtt gcattcgatt cctgtttgta attgtccttt      1800

taacagcgat cgcgtatttc gtctcgctca ggcgcaatca cgaatgaata cggtttggt      1860

tgatgcgagt gattttgatg acgagcgtaa tggctggcct gttgaacaag tctggaaaga      1920

aatgcataaa cttttgccat tctcaccgga ttcagtcgtc actcatggtg atttctcact      1980

tgataacctt attttgacg aggggaaatt aataggttgt attgatgttg acgagtcgg      2040

aatcgcagac cgataccagg atcttgccat cctatggaac tgcctcggtg agttttctcc      2100

ttcattacag aaacggcttt ttcaaaaata tggtattgat aatcctgata tgaataaatt      2160

gcagtttcat ttgatgctcg atgagttttt ctaatcagaa ttggttaatt ggttgtaaca      2220

ctggcagagc attacgctga cttgacggga cggcgcaagc tcatgaccaa atcccttaa      2280

cgtgagttac gcgtcgttcc actgagcgtc agaccccgta gaaaagatca aaggatcttc      2340

ttgagatcct ttttttctgc gcgtaatctg ctgcttgcaa acaaaaaaac caccgctacc      2400

agcggtggtt tgtttgccgg atcaagagct accaactctt tttccgaagg taactggctt      2460

cagcagagcg cagataccaa atactgtcct tctagtgtag ccgtagttag gccaccactt      2520

caagaactct gtagcaccgc ctacatacct cgctctgcta atcctgttac cagtggctgc      2580

tgccagtggc gataagtcgt gtcttaccgg gttggactca agacgatagt taccggataa      2640

ggcgcagcgg tcgggctgaa cggggggttc gtgcacacag cccagcttgg agcgaacgac      2700

ctacaccgaa ctgagatacc tacagcgtga gcattgagaa agcgccacgc ttcccgaagg      2760

gagaaaggcg acaggtatc cggtaagcgg cagggtcgga acaggagagc gcacagggga      2820

gcttccaggg ggaaacgcct ggtatcttta tagtcctgtc gggtttcgcc acctctgact      2880

tgagcgtcga tttttgtgat gctcgtcagg ggggcggagc ctatggaaaa acgccagcaa      2940

cgcggccttt ttacggttcc tggccttttg ctggcctttt gctcacatgt t      2991
```

```
<210> SEQ ID NO 39
<211> LENGTH: 13807
<212> TYPE: DNA
<213> ORGANISM: artificial
<220> FEATURE:
<223> OTHER INFORMATION: vector

<400> SEQUENCE: 39
```

```
aagctggtac gattgtaata cgactcacta tagggcgaat tgagcgctgt ttaaacgctc        60

ttcaactgga agagcggtta ccagagctgg tcacctttgt ccaccaagat ggaactgcgg       120

ccgctcatta attaagtcag gcgcgcctct agttgaagac acgttcatgt cttcatcgta       180

agaagacact cagtagtctt cggccagaat ggccgtaggt gaattaagag gagagaggag       240
```

-continued

```
gtaaacattt tcttctattt tttcatattt tcaggataaa ttattgtaaa agtttacaag    300

atttccattt gactagtgta aatgaggaat attctctagt aagatcatta tttcatctac    360

ttcttttatc ttctaccagt agaggaataa acaatattta gctcctttgt aaatacaaat    420

taattttcgt tcttgacatc attcaatttt aattttacgt ataaaataaa agatcatacc    480

tattagaacg attaaggaga aatacaattc gaatgagaag gatgtgccgt ttgttataat    540

aaacagccac acgacgtaaa cgtaaaatga ccacatgatg ggccaataga catggaccga    600

ctactaataa tagtaagtta cattttagga tggaataaat atcataccga catcagtttg    660

aaagaaaagg gaaaaaaaga aaaaataaat aaaagatata ctaccgacat gagttccaaa    720

aagcaaaaaa aaagatcaag ccgacacaga cacgcgtaga gagcaaaatg actttgacgt    780

cacaccacga aaacagacgc ttcatacgtg tccctttatc tctctcagtc tctctataaa    840

cttagtgaga ccctcctctg ttttactcag gatccccggg taccgagctc gaattcaccg    900

gtcgccacca tggcccacag caagcacggc ctgaaggagg agatgaccat gaagtaccac    960

atggagggct gcgtgaacgg ccacaagttc gtgatcaccg gcgagggcat cggctacccc   1020

ttcaagggca gcagaccat caacctgtgc gtgatcgagg gcggcccct gcccttcagc   1080

gaggacatcc tgagcgccgg cttcaagtac ggcgaccgga tcttcaccga gtaccccag   1140

gacatcgtgg actacttcaa gaacagctgc cccgccggct acacctgggg ccggagcttc   1200

ctgttcgagg acggcgccgt gtgcatctgt aacgtggaca tcaccgtgag cgtgaaggag   1260

aactgcatct accacaagag catcttcaac ggcgtgaact tccccgccga cggccccgtg   1320

atgaagaaga tgaccaccaa ctgggaggcc agctgcgaga agatcatgcc cgtgcctaag   1380

cagggcatcc tgaagggcga cgtgagcatg tacctgctgc tgaaggacgg cggccggtac   1440

cggtgccagt tcgacaccgt gtacaaggcc aagagcgtgc ccagcaagat gcccgagtgg   1500

cacttcatcc agcacaagct gctgcgggag gaccggagcg acgccaagaa ccagaagtgg   1560

cagctgaccg agcacgccat cgccttcccc agcgccctgg cctgaagcgg cccatggata   1620

ttcgaacgcg taggtaccac atggttaacc tagacttgtc catcttctgg attggccaac   1680

ttaattaatg tatgaaataa aaggatgcac acatagtgac atgctaatca ctataatgtg   1740

ggcatcaaag ttgtgtgtta tgtgtaatta ctagttatct gaataaaaga gaaagagatc   1800

atccatattt cttatcctaa atgaatgtca cgtgtcttta taattctttg atgaaccaga   1860

tgcatttcat taaccaaatc catatacata taaatattaa tcatatataa ttaatatcaa   1920

ttgggttagc aaaacaaatc tagtctaggt gtgttttgcg aatgcggcca ttggcctaga   1980

aggccatttta aatcctgagg atctggtctt cctaaggacc cgggatatcg ctatcaactt   2040

tgtatagaaa agttgaacga aaacgtaaa atgatataaa tatcaatata ttaaattaga   2100

ttttgcataa aaaacagact acataatact gtaaaacaca acatatccag tcactatggt   2160

cgacctgcag actggctgtg tataagggag cctgacattt atattcccca gaacatcagg   2220

ttaatggcgt ttttgatgtc attttcgcgg tggctgagat cagccacttc ttccccgata   2280

acggagaccg gcacactggc catatcggtg tcatcatgc gccagctttc atccccgata   2340

tgcaccaccg ggtaaagttc acgggggact ttatctgaca gcagacgtgc actggccagg   2400

gggatcacca tccgtcgccc gggcgtgtca ataatatcac tctgtacatc cacaaacaga   2460

cgataacggc tctctctttt ataggtgtaa accttaaact gcatttcacc agcccctgtt   2520
```

-continued

```
ctcgtcggca aaagagccgt tcatttcaat aaaccgggcg acctcagcca tcccttcctg    2580

attttccgct ttccagcgtt cggcacgcag acgacgggct tcattctgca tggttgtgct    2640

taccgaaccg agatattga catcatatat gccttgagca actgatagct gtcgctgtca     2700

actgtcactg taatacgctg cttcatagca tacctctttt tgacatactt cgggtataca    2760

tatcagtata tattcttata ccgcaaaaat cagcgcgcaa atacgcatac tgttatctgg    2820

cttttagtaa gccggatcct ctagattacg ccccgcctgc cactcatcgc agtactgttg    2880

taattcatta agcattctgc cgacatggaa gccatcacaa acggcatgat gaacctgaat    2940

cgccagcggc atcagcacct tgtcgccttg cgtataatat ttgcccatgg tgaaaacggg    3000

ggcgaagaag ttgtccatat tggccacgtt taaatcaaaa ctggtgaaac tcacccaggg    3060

attggctgag acgaaaaaca tattctcaat aaacccttta gggaaatagg ccaggttttc    3120

accgtaacac gccacatctt gcgaatatat gtgtagaaac tgccggaaat cgtcgtggta    3180

ttcactccag agcgatgaaa acgtttcagt ttgctcatgg aaaacggtgt aacaagggtg    3240

aacactatcc catatcacca gctcaccgtc tttcattgcc atacggaatt ccggatgagc    3300

attcatcagg cgggcaagaa tgtgaataaa ggccggataa aacttgtgct tatttttctt    3360

tacggtcttt aaaaaggccg taatatccag ctgaacggtc tggttatagg tacattgagc    3420

aactgactga atgcctcaa aatgttcttt acgatgccat gggatatat caacggtggt      3480

atatccagtg atttttttct ccattttagc ttccttagct cctgaaaatc tcgacggatc    3540

ctaactcaaa atccacacat tatacgagcc ggaagcataa agtgtaaagc ctggggtgcc    3600

ctaatgcggc cgccatagtg actggatatg ttgtgtttta cagtattatg tagtctgttt    3660

tttatgcaaa atctaattta atatattgat atttatatca ttttacgttt ctcgttcaac    3720

tttattatac aaagttgata gatatcggac cgattaaact ttaattcggt ccgaagcttg    3780

catgcctgca gtgcagcgtg acccggtcgt gccctctct agagataatg agcattgcat      3840

gtctaagtta taaaaaatta ccacatattt tttttgtcac acttgtttga agtgcagttt    3900

atctatcttt atacatatat ttaaactta ctctacgaat aatataatct atagtactac      3960

aataatatca gtgtttaga gaatcatata aatgaacagt tagacatggt ctaaaggaca      4020

attgagtatt ttgacaacag gactctacag ttttatcttt ttagtgtgca tgtgttctcc    4080

ttttttttg caaatagctt cacctatata atacttcatc cattttatta gtacatccat      4140

ttagggttta gggttaatgg ttttttataga ctaattttt tagtacatct attttattct     4200

attttagcct ctaaattaag aaaactaaaa ctctatttta gttttttat ttaataattt      4260

agatataaaa tagaataaaa taaagtgact aaaaattaaa caaataccct ttaagaaatt    4320

aaaaaaacta aggaaacatt tttcttgttt cgagtagata atgccagcct gttaaacgcc    4380

gtcgacgagt ctaacggaca ccaaccagcg aaccagcagc gtcgcgtcgg ccaagcgaa     4440

gcagacggca cggcatctct gtcgctgcct ctggacccct ctcgagagtt ccgctccacc    4500

gttggacttg ctccgctgtc ggcatccaga aattgcgtgg cggagcggca gacgtgagcc    4560

ggcacggcag cgggcctcct cctcctctca cggcaccggc agctacgggg gattcctttc    4620

ccaccgctcc ttcgctttcc cttcctcgcc cgccgtaata aatagacacc ccctccacac    4680

cctctttccc caacctcgtg ttgttcggag cgcacacaca cacaaccaga tctcccccaa    4740

atccacccgt cggcacctcc gcttcaaggt acgccgctcg tcctcccccc cccccctctc    4800
```

-continued

```
taccttctct agatcggcgt tccggtccat gcatggttag ggcccggtag ttctacttct    4860

gttcatgttt gtgttagatc cgtgtttgtg ttagatccgt gctgctagcg ttcgtacacg    4920

gatgcgacct gtacgtcaga cacgttctga ttgctaactt gccagtgttt ctctttgggg    4980

aatcctggga tggctctagc cgttccgcag acgggatcga tttcatgatt ttttttgttt    5040

cgttgcatag ggtttggttt gcccttttcc tttatttcaa tatatgccgt gcacttgttt    5100

gtcgggtcat cttttcatgc tttttttttgt cttggttgtg atgatgtggt ctggttgggc    5160

ggtcgttcta gatcggagta gaattctgtt tcaaactacc tggtggattt attaattttg    5220

gatctgtatg tgtgtgccat acatattcat agttacgaat tgaagatgat ggatggaaat    5280

atcgatctag gataggtata catgttgatg cgggtttttac tgatgcatat acagagatgc    5340

tttttgttcg cttggttgtg atgatgtggt gtggttgggc ggtcgttcat tcgttctaga    5400

tcggagtaga atactgtttc aaactacctg gtgtatttat taattttgga actgtatgtg    5460

tgtgtcatac atcttcatag ttacgagttt aagatggatg gaaatatcga tctaggatag    5520

gtatacatgt tgatgtgggt tttactgatg catatacatg atggcatatg cagcatctat    5580

tcatatgctc taaccttgag tacctatcta ttataataaa caagtatgtt ttataattat    5640

tttgatcttg atatacttgg atgatggcat atgcagcagc tatatgtgga ttttttttagc    5700

cctgccttca tacgctattt atttgcttgg tactgtttct tttgtcgatg ctcaccctgt    5760

tgtttggtgt tacttctgca ggtcgacttt aacttagcct aggatccaca cgacaccatg    5820

tcccccgagc gccgcccccgt cgagatccgc ccggccaccg ccgccgacat ggccgccgtg    5880

tgcgacatcg tgaaccacta catcgagacc tccaccgtga acttccgcac cgagccgcag    5940

accccgcagg agtggatcga cgacctggag cgcctccagg accgctaccc gtggctcgtg    6000

gccgaggtgg agggcgtggt ggccggcatc gcctacgccg cccgtggaa ggcccgcaac    6060

gcctacgact ggaccgtgga gtccaccgtg tacgtgtccc accgccacca gcgcctcggc    6120

ctcggctcca ccctctacac ccacctcctc aagagcatgg aggcccaggg cttcaagtcc    6180

gtggtggccg tgatcggcct cccgaacgac ccgtccgtgc gcctccacga ggccctcggc    6240

tacaccgccc gcggcacccct ccgcgccgcc ggctacaagc acggcggctg gcacgacgtc    6300

ggcttctggc agcgcgactt cgagctgccg gccccgccgc gcccggtgcg cccggtgacg    6360

cagatctccg gtggaggcgg cagcggtggc ggaggctccg gaggcggtgg ctccatggcc    6420

tcctccgagg acgtcatcaa ggagttcatg cgcttcaagg tgcgcatgga gggctccgtg    6480

aacggccacg agttcgagat cgagggcgag ggcgagggcc gcccctacga gggcacccag    6540

accgccaagc tgaaggtgac caagggcggc cccctgcccct tcgcctggga catcctgtcc    6600

ccccagttcc agtacggctc caaggtgtac gtgaagcacc ccgccgacat ccccgactac    6660

aagaagctgt ccttccccga gggcttcaag tgggagcgcg tgatgaactt cgaggacggc    6720

ggcgtggtga ccgtgaccca ggactcctcc ctgcaggacg gctccttcat ctacaaggtg    6780

aagttcatcg gcgtgaactt ccccctccgac ggccccgtaa tgcagaagaa gactatgggc    6840

tgggaggcct ccaccgagcg cctgtacccc cgcgacggcg tgctgaaggg cgagatccac    6900

aaggccctga gctgaagga cggcggccac tacctggtgg agttcaagtc catctacatg    6960

gccaagaagc ccgtgcagct gcccggctac tactacgtgg actccaagct ggacatcacc    7020

tcccacaacg aggactacac catcgtggag cagtacgagc gcgccgaggg ccgccaccac    7080
```

-continued

```
ctgttcctgt agtcaggatc tgagtcgaaa cctagacttg tccatcttct ggattggcca      7140

acttaattaa tgtatgaaat aaaaggatgc acacatagtg acatgctaat cactataatg      7200

tgggcatcaa agttgtgtgt tatgtgtaat tactagttat ctgaataaaa gagaaagaga      7260

tcatccatat ttcttatcct aaatgaatgt cacgtgtctt tataattctt tgatgaacca      7320

gatgcatttc attaaccaaa tccatataca tataaatatt aatcatatat aattaatatc      7380

aattgggtta gcaaaacaaa tctagtctag gtgtgttttg cgaatgcggc cgccaccgcg      7440

gtggagctcg aattcattcc gattaatcgt ggcctcttgc tcttcaggat gaagagctat      7500

gtttaaacgt gcaagcgcta ctagacaatt cagtacatta aaaacgtccg caatgtgtta      7560

ttaagttgtc taagcgtcaa tttgtttaca ccacaatata tcctgccacc agccagccaa      7620

cagctccccg accggcagct cggcacaaaa tcaccactcg atacaggcag cccatcagtc      7680

cgggacggcg tcagcgggag agccgttgta aggcggcaga ctttgctcat gttaccgatg      7740

ctattcggaa gaacggcaac taagctgccg ggtttgaaac acggatgatc tcgcggaggg      7800

tagcatgttg attgtaacga tgacagagcg ttgctgcctg tgatcaaata tcatctccct      7860

cgcagagatc cgaattatca gccttcttat tcatttctcg cttaaccgtg acaggctgtc      7920

gatcttgaga actatgccga cataatagga aatcgctgga taaagccgct gaggaagctg      7980

agtggcgcta tttcttttaga agtgaacgtt gacgatcgtc gaccgtaccc cgatgaatta      8040

attcggacgt acgttctgaa cacagctgga tacttacttg ggcgattgtc atacatgaca      8100

tcaacaatgt acccgtttgt gtaaccgtct cttggaggtt cgtatgcac tagtggttcc      8160

cctcagcttg cgactagatg ttgaggccta acattttatt agagagcagg ctagttgctt      8220

agatacatga tcttcaggcc gttatctgtc agggcaagcg aaaattggcc atttatgacg      8280

accaatgccc cgcagaagct cccatctttg ccgccataga cgccgcgccc cccttttggg      8340

gtgtagaaca tccttttgcc agatgtggaa aagaagttcg ttgtcccatt gttggcaatg      8400

acgtagtagc cggcgaaagt gcgagaccca tttgcgctat atataagcct acgatttccg      8460

ttgcgactat tgtcgtaatt ggatgaacta ttatcgtagt tgctctcaga gttgtcgtaa      8520

tttgatggac tattgtcgta attgcttatg gagttgtcgt agttgcttgg agaaatgtcg      8580

tagttggatg gggagtagtc atagggaaga cgagcttcat ccactaaaac aattggcagg      8640

tcagcaagtg cctgccccga tgccatcgca agtacgaggc ttagaaccac cttcaacaga      8700

tcgcgcatag tcttccccag ctctctaacg cttgagttaa gccgcgccgc gaagcggcgt      8760

cggcttgaac gaattgttag acattatttg ccgactacct tggtgatctc gcctttcacg      8820

tagtgaacaa attcttccaa ctgatctgcg cgcgaggcca agcgatcttc ttgtccaaga      8880

taagcctgcc tagcttcaag tatgacgggc tgatactggg ccggcaggcg ctccattgcc      8940

cagtcggcag cgacatcctt cggcgcgatt ttgccggtta ctgcgctgta ccaaatgcgg      9000

gacaacgtaa gcactacatt tcgctcatcg ccagcccagt cgggcggcga gttccatagc      9060

gttaaggttt catttagcgc ctcaaataga tcctgttcag gaaccggatc aaagagttcc      9120

tccgccgctg gacctaccaa ggcaacgcta tgttctcttg cttttgtcag caagatagcc      9180

agatcaatgt cgatcgtggc tggctcgaag atacctgcaa gaatgtcatt gcgctgccat      9240

tctccaaatt gcagttcgcg cttagctgga taacgccacg gaatgatgtc gtcgtgcaca      9300

acaatggtga cttctacagc gcggagaatc tcgctctctc caggggaagc cgaagtttcc      9360
```

-continued

```
aaaaggtcgt tgatcaaagc tcgccgcgtt gtttcatcaa gccttacagt caccgtaacc   9420

agcaaatcaa tatcactgtg tggcttcagg ccgccatcca ctgcggagcc gtacaaatgt   9480

acggccagca acgtcggttc gagatggcgc tcgatgacgc caactacctc tgatagttga   9540

gtcgatactt cggcgatcac cgcttccctc atgatgttta actcctgaat taagccgcgc   9600

cgcgaagcgg tgtcggcttg aatgaattgt taggcgtcat cctgtgctcc cgagaaccag   9660

taccagtaca tcgctgtttc gttcgagact tgaggtctag ttttatacgt gaacaggtca   9720

atgccgccga gagtaaagcc acattttgcg tacaaattgc aggcaggtac attgttcgtt   9780

tgtgtctcta atcgtatgcc aaggagctgt ctgcttagtg cccacttttt cgcaaattcg   9840

atgagactgt gcgcgactcc tttgcctcgg tgcgtgtgcg acacaacaat gtgttcgata   9900

gaggctagat cgttccatgt tgagttgagt tcaatcttcc cgacaagctc ttggtcgatg   9960

aatgcgccat agcaagcaga gtcttcatca gagtcatcat ccgagatgta atccttccgg   10020

taggggctca cacttctggt agatagttca aagccttggt cggataggtg cacatcgaac   10080

acttcacgaa caatgaaatg gttctcagca tccaatgttt ccgccacctg ctcagggatc   10140

accgaaatct tcatatgacg cctaacgcct ggcacagcgg atcgcaaacc tggcgcggct   10200

tttggcacaa aaggcgtgac aggtttgcga atccgttgct gccacttgtt aaccctttg   10260

ccagatttgg taactataat ttatgttaga ggcgaagtct tgggtaaaaa ctggcctaaa   10320

attgctgggg atttcaggaa agtaaacatc accttccggc tcgatgtcta ttgtagatat   10380

atgtagtgta tctacttgat cggggggatct gctgcctcgc gcgtttcggt gatgacggtg   10440

aaaacctctg acacatgcag ctcccggaga cggtcacagc ttgtctgtaa gcggatgccg   10500

ggagcagaca agcccgtcag ggcgcgtcag cgggtgttgg cgggtgtcgg ggcgcagcca   10560

tgacccagtc acgtagcgat agcggagtgt atactggctt aactatgcgg catcagagca   10620

gattgtactg agagtgcacc atatgcggtg tgaaataccg cacagatgcg taaggagaaa   10680

ataccgcatc aggcgctctt ccgcttcctc gctcactgac tcgctgcgct cggtcgttcg   10740

gctgcggcga gcggtatcag ctcactcaaa ggcggtaata cggttatcca cagaatcagg   10800

ggataacgca ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa   10860

ggccgcgttg ctggcgtttt tccataggct ccgcccccct gacgagcatc acaaaaatcg   10920

acgctcaagt cagaggtggc gaaacccgac aggactataa agataccagg cgtttccccc   10980

tggaagctcc ctcgtgcgct ctcctgttcc gaccctgccg cttaccggat acctgtccgc   11040

ctttctccct tcgggaagcg tggcgctttc tcatagctca cgctgtaggt atctcagttc   11100

ggtgtaggtc gttcgctcca agctgggctg tgtgcacgaa ccccccgttc agcccgaccg   11160

ctgcgcctta tccggtaact atcgtcttga gtccaacccg gtaagacacg acttatcgcc   11220

actggcagca gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga   11280

gttcttgaag tggtggccta actacggcta cactagaagg acagtatttg gtatctgcgc   11340

tctgctgaag ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaacaaac   11400

caccgctggt agcggtggtt tttttgtttg caagcagcag attacgcgca gaaaaaaagg   11460

atctcaagaa gatcctttga tcttttctac ggggtctgac gctcagtgga cgaaaactc   11520

acgttaaggg attttggtca tgagattatc aaaaaggatc ttcacctaga tcctttaaaa   11580

ttaaaaatga agttttaaat caatctaaag tatatatgag taaacttggt ctgacagtta   11640
```

-continued

```
ccaatgctta atcagtgagg cacctatctc agcgatctgt ctatttcgtt catccatagt    11700

tgcctgactc cccgtcgtgt agataactac gatacgggag ggcttaccat ctggccccag    11760

tgctgcaatg ataccgcgag acccacgctc accggctcca gatttatcag caataaacca    11820

gccagccgga agggccgagc gcagaagtgg tcctgcaact ttatccgcct ccatccagtc    11880

tattaattgt tgccgggaag ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt    11940

tgttgccatt gctgcagggg gggggggggg ggggacttc cattgttcat tccacggaca    12000

aaaacagaga aaggaaacga cagaggccaa aaagcctcgc tttcagcacc tgtcgtttcc    12060

tttcttttca gagggtattt taaataaaaa cattaagtta tgacgaagaa gaacggaaac    12120

gccttaaacc ggaaaatttt cataaatagc gaaaacccgc gaggtcgccg ccccgtaacc    12180

tgtcggatca ccggaaagga cccgtaaagt gataatgatt atcatctaca tatcacaacg    12240

tgcgtggagg ccatcaaacc acgtcaaata atcaattatg acgcaggtat cgtattaatt    12300

gatctgcatc aacttaacgt aaaaacaact tcagacaata caaatcagcg acactgaata    12360

cggggcaacc tcatgtcccc ccccccccc ccctgcagg catcgtggtg tcacgctcgt    12420

cgtttggtat ggcttcattc agctccggtt cccaacgatc aaggcgagtt acatgatccc    12480

ccatgttgtg caaaaaagcg gttagctcct tcggtcctcc gatcgttgtc agaagtaagt    12540

tggccgcagt gttatcactc atggttatgg cagcactgca taattctctt actgtcatgc    12600

catccgtaag atgcttttct gtgactggtg agtactcaac caagtcattc tgagaatagt    12660

gtatgcggcg accgagttgc tcttgcccgg cgtcaacacg ggataatacc gcgccacata    12720

gcagaacttt aaaagtgctc atcattggaa aacgttcttc gggcgaaaa ctctcaagga    12780

tcttaccgct gttgagatcc agttcgatgt aacccactcg tgcacccaac tgatcttcag    12840

catcttttac tttcaccagc gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa    12900

aaaagggaat aagggcgaca cggaaatgtt gaatactcat actcttcctt tttcaatatt    12960

attgaagcat ttatcagggt tattgtctca tgagcggata catatttgaa tgtatttaga    13020

aaaataaaca aataggggtt ccgcgcacat ttccccgaaa agtgccacct gacgtctaag    13080

aaaccattat tatcatgaca ttaacctata aaaataggcg tatcacgagg ccctttcgtc    13140

ttcaagaatt ggtcgacgat cttgctgcgt tcggatattt tcgtggagtt cccgccacag    13200

acccggattg aaggcgagat ccagcaactc gcgccagatc atcctgtgac ggaactttgg    13260

cgcgtgatga ctggccagga cgtcggccga aagagcgaca agcagatcac gcttttcgac    13320

agcgtcggat ttgcgatcga ggattttttcg gcgctgcgct acgtccgcga ccgcgttgag    13380

ggatcaagcc acagcagccc actcgacctt ctagccgacc cagacgagcc aagggatctt    13440

tttggaatgc tgctccgtcg tcaggctttc cgacgtttgg gtggttgaac agaagtcatt    13500

atcgtacgga atgccaagca ctcccgaggg gaaccctgtg gttggcatgc acatacaaat    13560

ggacgaacgg ataaaccttt tcacgccctt ttaaatatcc gttattctaa taaacgctct    13620

tttctcttag gtttacccgc caatatatcc tgtcaaacac tgatagttta aactgaaggc    13680

gggaaacgac aatctgatca tgagcggaga attaagggag tcacgttatg accccccgccg    13740

atgacgcggg acaagccgtt ttacgtttgg aactgacaga accgcaacgt tgaaggagcc    13800

actcagc                                                             13807
```

-continued

<210> SEQ ID NO 40
<211> LENGTH: 3505
<212> TYPE: DNA
<213> ORGANISM: artificial
<220> FEATURE:
<223> OTHER INFORMATION: vector

<400> SEQUENCE: 40

```
gatccccggg taccgagctc gaattcggcc caagtttgta caaaaaagtt gaacgagaaa      60

cgtaaaatga tataaatatc aatatattaa attagatttt gcataaaaaa cagactacat     120

aatactgtaa aacacaacat atgcagtcac tatgaatcaa ctacttagat ggtattagtg     180

acctgtagaa ttcgagctct agagctgcag ggcggccgcg atatcccta tagtgagtcg     240

tattacatgg tcatagctgt ttcctggcag ctctggcccg tgtctcaaaa tctctgatgt     300

tacattgcac aagataaaaa tatatcatca tgaacaataa aactgtctgc ttacataaac     360

agtaatacaa ggggtgttat gagccatatt caacgggaaa cgtcgaggcc gcgattaaat     420

tccaacatgg atgctgattt atatgggtat aaatgggctc gcgataatgt cgggcaatca     480

ggtgcgacaa tctatcgctt gtatgggaag cccgatgcgc cagagttgtt tctgaaacat     540

ggcaaaggta gcgttgccaa tgatgttaca gatgagatgg tcagactaaa ctggctgacg     600

gaatttatgc ctcttccgac catcaagcat tttatccgta ctcctgatga tgcatggtta     660

ctcaccactg cgatccccgg aaaaacagca ttccaggtat tagaagaata tcctgattca     720

ggtgaaaata ttgttgatgc gctggcagtg ttcctgcgcc ggttgcattc gattcctgtt     780

tgtaattgtc cttttaacag cgatcgcgta tttcgtctcg ctcaggcgca atcacgaatg     840

aataacggtt tggttgatgc gagtgatttt gatgacgagc gtaatggctg gcctgttgaa     900

caagtctgga agaaatgca taaacttttg ccattctcac cggattcagt cgtcactcat     960

ggtgatttct cacttgataa ccttattttt gacgagggga attaatagg ttgtattgat    1020

gttggacgag tcggaatcgc agaccgatac caggatcttg ccatcctatg gaactgcctc    1080

ggtgagtttt ctccttcatt acagaaacgg ctttttcaaa aatatggtat tgataatcct    1140

gatatgaata aattgcagtt tcatttgatg ctcgatgagt ttttctaatc agaattggtt    1200

aattggttgt aacactggca gagcattacg ctgacttgac gggacggcgc aagctcatga    1260

ccaaaatccc ttaacgtgag ttacgcgtcg ttccactgag cgtcagaccc cgtagaaaag    1320

atcaaaggat cttcttgaga tccttttttt ctgcgcgtaa tctgctgctt gcaaacaaaa    1380

aaaccaccgc taccagcggt ggtttgtttg ccggatcaag agctaccaac tctttttccg    1440

aaggtaactg gcttcagcag agcgcagata ccaaatactg tccttctagt gtagccgtag    1500

ttaggccacc acttcaagaa ctctgtagca ccgcctacat acctcgctct gctaatcctg    1560

ttaccagtgg ctgctgccag tggcgataag tcgtgtctta ccgggttgga ctcaagacga    1620

tagttaccgg ataaggcgca gcggtcgggc tgaacggggg gttcgtgcac acagcccagc    1680

ttggagcgaa cgacctacac cgaactgaga tacctacagc gtgagcattg agaaagcgcc    1740

acgcttcccg aagggagaaa ggcggacagg tatccggtaa gcggcagggt cggaacagga    1800

gagcgcacga gggagcttcc agggggaaac gcctggtatc tttatagtcc tgtcgggttt    1860

cgccacctct gacttgagcg tcgatttttg tgatgctcgt caggggggcg gagcctatgg    1920

aaaaacgcca gcaacgcggc cttttttacgg ttcctggcct tttgctggcc ttttgctcac    1980

atgttctttc ctgcgttatc ccctgattct gtggataacc gtattaccgc ctttgagtga    2040
```

-continued

```
gctgataccg ctcgccgcag ccgaacgacc gagcgcagcg agtcagtgag cgaggaagcg      2100

gaagagcgcc caatacgcaa accgcctctc cccgcgcgtt ggccgattca ttaatgcagc      2160

tggcacgaca ggtttcccga ctggaaagcg ggcagtgagc gcaacgcaat taatacgcgt      2220

accgctagcc aggaagagtt tgtagaaacg caaaaaggcc atccgtcagg atggccttct      2280

gcttagtttg atgcctggca gtttatggcg ggcgtcctgc ccgccaccct ccgggccgtt      2340

gcttcacaac gttcaaatcc gctcccggcg gatttgtcct actcaggaga gcgttcaccg      2400

acaaacaaca gataaaacga aaggcccagt cttccgactg agcctttcgt tttatttgat      2460

gcctggcagt tccctactct cgcgttaacg ctagcatgga gtttttccca gtcacgacgt      2520

tgtaaaacga cggccagtct taagctcggg cccgcgttaa cgctaccatg gagctccaaa      2580

taatgatttt attttgactg atagtgacct gttcgttgca acaaattgat aagcaatgct      2640

tttttataat gccaactttg tatagaaaag ttgaagctta aatccttaca gaattgctgt      2700

agtttcatag tgctagatgt ggacagcaaa gcgccgctgt atgcttctgc ttttcttttt      2760

tggtgtgtgt agccacatcc tttgttcctg cccggcgcca tcccacttgg ttgttttttt      2820

ttatgattga aagccttcat gcttcctcgg tcaatcaccg gtgcgcactg ggagcatcgc      2880

cggaaaaaaa attcttcggc taagagtaac ttctttctcc ttttcttctc tgatctcgcg      2940

agcagtgctg ataacgtgtt gtaatctact tagcggtaac gagattgaga gagacaaaat      3000

gacagaacta ttgtctttat tgcagagtgt catgtattta tacaggggat acaaagtctc      3060

ccaaggggtg tgtcccttgg gagtaactgc cagttgatca caggacaata ttttgtaaca      3120

aaacgtacac atcgtcaaaa tagcgaggca tgaaactggc cttggccatg gacgcgtgaa      3180

gcgcgccatg cgttggatat gtggtcaata agtatataca atacaatgtt taacagagct      3240

gatagtactg ctttggcaca ttttttgtcca cgcttcatga gagataaaac acctgcacgt      3300

aaattcacat gctgcactga aggcccgatc actgaggagc gaactgccgt aactcccttc      3360

tatatatacc cccagtccct gtttcagttt tcgtcaagct agcagcacca agttgtcgat      3420

cacttgcctg ctcttgagct cgattaagct atcatcagct acagcatccg atcccaaact      3480

gcaactgtag cagcgacaac tgccg                                           3505
```

<210> SEQ ID NO 41
<211> LENGTH: 4778
<212> TYPE: DNA
<213> ORGANISM: artificial
<220> FEATURE:
<223> OTHER INFORMATION: vector

<400> SEQUENCE: 41

```
tggctgacgg aatttatgcc tcttccgacc atcaagcatt ttatccgtac tcctgatgat       60

gcatggttac tcaccactgc gatccccgga aaaacagcat tccaggtatt agaagaatat      120

cctgattcag gtgaaaatat tgttgatgcg ctggcagtgt tcctgcgccg gttgcattcg      180

attcctgttt gtaattgtcc ttttaacagc gatcgcgtat ttcgtctcgc tcaggcgcaa      240

tcacgaatga ataacggttt ggttgatgcg agtgattttg atgacgagcg taatggctgg      300

cctgttgaac aagtctggaa agaaatgcat aaacttttgc cattctcacc ggattcagtc      360

gtcactcatg gtgatttctc acttgataac cttatttttg acgaggggaa attaataggt      420

tgtattgatg ttggacgagt cggaatcgca gaccgatacc aggatcttgc catcctatgg      480
```

-continued

```
aactgcctcg gtgagttttc tccttcatta cagaaacggc ttttttcaaaa atatggtatt      540

gataatcctg atatgaataa attgcagttt catttgatgc tcgatgagtt tttctaatca      600

gaattggtta attggttgta acactggcag agcattacgc tgacttgacg ggacggcgca      660

agctcatgac caaaatccct taacgtgagt tacgcgtcgt tccactgagc gtcagacccc      720

gtagaaaaga tcaaaggatc ttcttgagat cctttttttc tgcgcgtaat ctgctgcttg      780

caaacaaaaa aaccaccgct accagcggtg gtttgtttgc cggatcaaga gctaccaact      840

cttttttccga aggtaactgg cttcagcaga gcgcagatac caaatactgt ccttctagtg      900

tagccgtagt taggccacca cttcaagaac tctgtagcac cgcctacata cctcgctctg      960

ctaatcctgt taccagtggc tgctgccagt ggcgataagt cgtgtcttac cgggttggac     1020

tcaagacgat agttaccgga taaggcgcag cggtcgggct gaacggggggg ttcgtgcaca     1080

cagcccagct tggagcgaac gacctacacc gaactgagat acctacagcg tgagcattga     1140

gaaagcgcca cgcttcccga agggagaaag gcggacaggt atccggtaag cggcagggtc     1200

ggaacaggag agcgcacgag ggagcttcca gggggaaacg cctggtatct ttatagtcct     1260

gtcgggtttc gccacctctg acttgagcgt cgatttttgt gatgctcgtc aggggggcgg     1320

agcctatgga aaaacgccag caacgcggcc ttttttacggt tcctggcctt ttgctggcct     1380

tttgctcaca tgttctttcc tgcgttatcc cctgattctg tggataaccg tattaccgcc     1440

tttgagtgag ctgataccgc tcgccgcagc cgaacgaccg agcgcagcga tcagtgagc     1500

gaggaagcgg aagagcgccc aatacgcaaa ccgcctctcc ccgcgcgttg gccgattcat     1560

taatgcagct ggcacgacag gtttcccgac tggaaagcgg gcagtgagcg caacgcaatt     1620

aatacgcgta ccgctagcca ggaagagttt gtagaaacgc aaaaaggcca tccgtcagga     1680

tggccttctg cttagtttga tgcctggcag tttatggcgg gcgtcctgcc cgccaccctc     1740

cgggccgttg cttcacaacg ttcaaatccg ctcccggcgg atttgtccta ctcaggagag     1800

cgttcaccga caaacaacag ataaaacgaa aggcccagtc ttccgactga gcctttcgtt     1860

ttatttgatg cctggcagtt ccctactctc gcgttaacgc tagcatggat gttttcccag     1920

tcacgacgtt gtaaaacgac ggccagtctt aagctcgggc ccgcgttaac gctaccatgg     1980

agctccaaat aatgatttta ttttgactga tagtgacctg ttcgttgcaa caaattgata     2040

agcaatgctt ttttataatg ccaactttgt atagaaaagt tgggccgaat tcgagctcgg     2100

tacggccaga atggcccgga ccgggttacc gaattcgagc tcggtaccct gggatcagct     2160

tgcatgcctg cagtgcagcg tgaccggtc gtgcccctct ctagagataa tgagcattgc     2220

atgtctaagt tataaaaaat taccacatat ttttttttgtc acacttgttt gaagtgcagt     2280

ttatctatct ttatacatat atttaaactt tactctacga ataatataat ctatagtact     2340

acaataatat cagtgtttta gagaatcata taaatgaaca gttagacatg gtctaaagga     2400

caattgagta ttttgacaac aggactctac agtttttatct tttttagtgtg catgtgttct     2460

ccttttttttt tgcaaatagc ttcacctata taatacttca tccatttttat tagtacatcc     2520

atttagggtt tagggttaat ggttttttata gactaatttt tttagtacat ctattttatt     2580

ctattttagc ctctaaatta agaaaactaa aactctattt tagtttttttt atttaataat     2640

ttagatataa aatagaataa aataaagtga ctaaaaatta aacaaatacc ctttaagaaa     2700

ttaaaaaaac taaggaaaca tttttcttgt ttcgagtaga taatgccagc ctgttaaacg     2760
```

-continued

```
ccgtcgacga gtctaacgga caccaaccag cgaaccagca gcgtcgcgtc gggccaagcg    2820

aagcagacgg cacggcatct ctgtcgctgc ctctggaccc ctctcgagag ttccgctcca    2880

ccgttggact tgctccgctg tcggcatcca gaaattgcgt ggcggagcgg cagacgtgag    2940

ccggcacggc aggcggcctc ctcctcctct cacggcaccg gcagctacgg gggattcctt    3000

tcccaccgct ccttcgcttt cccttcctcg cccgccgtaa taaatagaca cccctccac    3060

accctctttc cccaacctcg tgttgttcgg agcgcacaca cacacaacca gatctccccc    3120

aaatccaccc gtcggcacct ccgcttcaag gtacgccgct cgtcctcccc cccccccctc    3180

tctaccttct ctagatcggc gttccggtcc atgcatggtt agggcccggt agttctactt    3240

ctgttcatgt ttgtgttaga tccgtgtttg tgttagatcc gtgctgctag cgttcgtaca    3300

cggatgcgac ctgtacgtca gacacgttct gattgctaac ttgccagtgt ttctctttgg    3360

ggaatcctgg gatggctcta gccgttccgc agacgggatc gatttcatga ttttttttgt    3420

ttcgttgcat agggtttggt ttgccctttt cctttatttc aatatatgcc gtgcacttgt    3480

ttgtcgggtc atctttttcat gctttttttt gtcttggttg tgatgatgtg gtctggttgg    3540

gcggtcgttc tagatcggag tagaattctg tttcaaacta cctggtggat ttattaattt    3600

tggatctgta tgtgtgtgcc atacatattc atagttacga attgaagatg atggatggaa    3660

atatcgatct aggataggta tacatgttga tgcgggtttt actgatgcat atacagagat    3720

gctttttgtt cgcttggttg tgatgatgtg gtgtggttgg gcggtcgttc attcgttcta    3780

gatcggagta gaatactgtt tcaaactacc tggtgtattt attaattttg gaactgtatg    3840

tgtgtgtcat acatcttcat agttacgagt ttaagatgga tggaaatatc gatctaggat    3900

aggtatacat gttgatgtgg gttttactga tgcatataca tgatggcata tgcagcatct    3960

attcatatgc tctaaccttg agtacctatc tattataata aacaagtatg ttttataatt    4020

attttgatct tgatatactt ggatgatggc atatgcagca gctatatgtg gatttttttta    4080

gccctgcctt catacgctat ttatttgctt ggtactgttt cttttgtcga tgctcaccct    4140

gttgtttggt gttacttctg caggtcgact ctagaggatc agcttggtca cccggtccgg    4200

gcctagaagg ccagcttcaa gtttgtacaa aaaagttgaa cgagaaacgt aaaatgatat    4260

aaatatcaat atattaaatt agattttgca taaaaaacag actacataat actgtaaaac    4320

acaacatatg cagtcactat gaatcaacta cttagatggt attagtgacc tgtagaattc    4380

gagctctaga gctgcagggc ggccgcgata tccctatag tgagtcgtat tacatggtca    4440

tagctgtttc ctgcagctc tggcccgtgt ctcaaaatct ctgatgttac attgcacaag    4500

ataaaaatat atcatcatga acaataaaac tgtctgctta cataaacagt aatacaaggg    4560

gtgttatgag ccatattcaa cgggaaacgt cgaggccgcg attaaattcc aacatggatg    4620

ctgattttata tgggtataaa tgggctcgcg ataatgtcgg gcaatcaggt gcgacaatct    4680

atcgcttgta tgggaagccc gatgcgccag agttgtttct gaaacatggc aaaggtagcg    4740

ttgccaatga tgttacagat gagatggtca gactaaac                            4778
```

<210> SEQ ID NO 42
<211> LENGTH: 13019
<212> TYPE: DNA
<213> ORGANISM: artificial
<220> FEATURE:
<223> OTHER INFORMATION: vector

-continued

<400> SEQUENCE: 42

```
gttacccgga ccgaagctta gcccgggcat gcctgcagtg cagcgtgacc cggtcgtgcc        60

cctctctaga gataatgagc attgcatgtc taagttataa aaaattacca catatttttt       120

ttgtcacact tgtttgaagt gcagtttatc tatctttata catatatta  aactttactc       180

tacgaataat ataatctata gtactacaat aatatcagtg ttttagagaa tcatataaat       240

gaacagttag acatggtcta aaggacaatt gagtattttg acaacaggac tctacagttt       300

tatctttta  gtgtgcatgt gttctccttt ttttttgcaa atagcttcac ctatataata       360

cttcatccat tttattagta catccattta gggtttaggg ttaatggttt ttatagacta       420

attttttag  tacatctatt ttattctatt ttagcctcta aattaagaaa actaaaactc       480

tattttagtt tttttattta ataatttaga tataaaatag aataaaataa agtgactaaa       540

aattaaacaa ataccctta  agaaattaaa aaaactaagg aaacattttt cttgtttcga       600

gtagataatg ccagcctgtt aaacgccgtc gacgagtcta acggacacca accagcgaac       660

cagcagcgtc gcgtcgggcc aagcgaagca gacggcacgg catctctgtc gctgcctctg       720

gacccctctc gagagttccg ctccaccgtt ggacttgctc cgctgtcggc atccagaaat       780

tgcgtggcgg agcggcagac gtgagccggc acggcaggcg gcctcctcct cctctcacgg       840

cacggcagct acgggggatt cctttcccac cgctccttcg ctttcccttc ctcgcccgcc       900

gtaataaata gacacccct  ccacccctc  tttccccaac ctcgtgttgt tcggagcgca       960

cacacacaca accagatctc ccccaaatcc accgtcggc  acctccgctt caaggtacgc      1020

cgctcgtcct cccccccccc ccctctctac cttctctaga tcggcgttcc ggtccatggt      1080

tagggcccgg tagttctact tctgttcatg tttgtgttag atccgtgttt gtgttagatc      1140

cgtgctgcta gcgttcgtac acggatgcga cctgtacgtc agacacgttc tgattgctaa      1200

cttgccagtg tttctctttg gggaatcctg ggatggctct agccgttccg cagacgggat      1260

cgatttcatg atttttttg  tttcgttgca tagggtttgg tttgcccttt tcctttatt       1320

caatatatgc cgtgcacttg tttgtcgggt catcttttca tgctttttt  tgtcttggtt      1380

gtgatgatgt ggtctggttg ggcggtcgtt ctagatcgga gtagaattct gtttcaaact      1440

acctggtgga tttattaatt ttggatctgt atgtgtgtgc catacatatt catagttacg      1500

aattgaagat gatggatgga aatatcgatc taggataggt atacatgttg atgcgggttt      1560

tactgatgca tatacagaga tgctttttgt tcgcttggtt gtgatgatgt ggtgtggttg      1620

ggcggtcgtt cattcgttct agatcggagt agaatactgt ttcaaactac ctggtgtatt      1680

tattaatttt ggaactgtat gtgtgtgtca tacatcttca tagttacgag tttaagatgg      1740

atggaaatat cgatctagga taggtataca tgttgatgtg ggtttactg  atgcatatac      1800

atgatggcat atgcagcatc tattcatatg ctctaacctt gagtacctat ctattataat      1860

aaacaagtat gttttataat tattttgatc ttgatatact ggatgatgg  catatgcagc      1920

agctatatgt ggatttttt  agccctgcct tcatacgcta tttatttgct tggtactgtt      1980

tcttttgtcg atgctcaccc tgttgtttgg tgttacttct gcaggtcgac tctagaggat      2040

ccacaagttt gtacaaaaaa gctgaacgag aaacgtaaaa tgatataaat atcaatatat      2100

taaattagat tttgcataaa aaacagacta cataatactg taaaacacaa catatccagt      2160

cactatggcg gccgcattag gcaccccagg ctttacactt tatgcttccg gctcgtataa      2220
```

-continued

```
tgtgtggatt ttgagttagg atttaaatac gcgttgatcc ggcttactaa aagccagata    2280

acagtatgcg tatttgcgcg ctgatttttg cggtataaga atatatactg atatgtatac    2340

ccgaagtatg tcaaaaagag gtatgctatg aagcagcgta ttacagtgac agttgacagc    2400

gacagctatc agttgctcaa ggcatatatg atgtcaatat ctccggtctg gtaagcacaa    2460

ccatgcagaa tgaagcccgt cgtctgcgtg ccgaacgctg gaaagcggaa aatcaggaag    2520

ggatggctga ggtcgcccgg tttattgaaa tgaacggctc ttttgctgac gagaacaggg    2580

gctggtgaaa tgcagtttaa ggtttacacc tataaaagag agagccgtta tcgtctgttt    2640

gtggatgtac agagtgatat cattgacacg cccggtcgac ggatggtgat ccccctggcc    2700

agtgcacgtc tgctgtcaga taaagtctcc cgtgaacttt acccggtggt gcatatcggg    2760

gatgaaagct ggcgcatgat gaccaccgat atggccagtg tgccggtctc cgttatcggg    2820

gaagaagtgg ctgatctcag ccaccgcgaa aatgacatca aaaacgccat taacctgatg    2880

ttctggggaa tataaatgtc aggctccctt atacacagcc agtctgcagg tcgaccatag    2940

tgactggata tgttgtgttt tacagtatta tgtagtctgt tttttatgca aaatctaatt    3000

taatatattg atatttatat cattttacgt ttctcgttca gctttcttgt acaaagtggt    3060

gttaacctag acttgtccat cttctggatt ggccaactta attaatgtat gaaataaaag    3120

gatgcacaca tagtgacatg ctaatcacta taatgtgggc atcaaagttg tgtgttatgt    3180

gtaattacta gttatctgaa taaaagagaa agagatcatc catatttctt atcctaaatg    3240

aatgtcacgt gtctttataa ttctttgatg aaccagatgc atttcattaa ccaaatccat    3300

atacatataa atattaatca tatataatta atatcaattg ggttagcaaa acaaatctag    3360

tctaggtgtg ttttgcgaat tgcggccgcc accgcggtgg agctcgaatt ccggtccggg    3420

tcacctttgt ccaccaagat ggaactgcgg ccgctcatta attaagtcag gcgcgcctct    3480

agttgaagac acgttcatgt cttcatcgta agaagacact cagtagtctt cggccagaat    3540

ggccatctgg attcagcagg cctagaaggc catttaaatc ctgaggatct ggtcttccta    3600

aggacccggg atatcggacc gattaaactt taattcggtc cgaagcttgc atgcctgcag    3660

tgcagcgtga cccggtcgtg cccctctcta gagataatga gcattgcatg tctaagttat    3720

aaaaaattac cacatatttt ttttgtcaca cttgtttgaa gtgcagttta tctatcttta    3780

tacatatatt taaactttac tctacgaata atataatcta tagtactaca ataatatcag    3840

tgttttagag aatcatataa atgaacagtt agacatggtc taaaggacaa ttgagtattt    3900

tgacaacagg actctacagt tttatctttt tagtgtgcat gtgttctcct tttttttttgc    3960

aaatagcttc acctatataa tacttcatcc attttattag tacatccatt tagggtttag    4020

ggttaatggt ttttatagac taattttttt agtacatcta ttttattcta ttttagcctc    4080

taaattaaga aaactaaaac tctattttag tttttttatt taataattta gatataaaat    4140

agaataaaat aaagtgacta aaaattaaac aaatacccct taagaaatta aaaaaactaa    4200

ggaaacattt ttcttgtttc gagtagataa tgccagcctg ttaaacgccg tcgacgagtc    4260

taacggacac caaccagcga accagcagcg tcgcgtcggg ccaagcgaag cagacggcac    4320

ggcatctctg tcgctgcctc tggacccctc tcgagagttc cgctccaccg ttggacttgc    4380

tccgctgtcg gcatccagaa attgcgtggc ggagcggcag acgtgagccg cacggcagg    4440

cggcctcctc ctcctctcac ggcaccggca gctacggggg attcctttcc caccgctcct    4500
```

-continued

```
tcgctttccc ttcctcgccc gccgtaataa atagacaccc cctccacacc ctctttcccc   4560

aacctcgtgt tgttcggagc gcacacacac acaaccagat ctcccccaaa tccacccgtc   4620

ggcacctccg cttcaaggta cgccgctcgt cctccccccc ccccctctct accttctcta   4680

gatcggcgtt ccggtccatg catggttagg gcccggtagt tctacttctg ttcatgtttg   4740

tgttagatcc gtgtttgtgt tagatccgtg ctgctagcgt tcgtacacgg atgcgacctg   4800

tacgtcagac acgttctgat tgctaacttg ccagtgtttc tctttgggga atcctgggat   4860

ggctctagcc gttccgcaga cgggatcgat ttcatgattt tttttgtttc gttgcatagg   4920

gtttggtttg cccttttcct ttatttcaat atatgccgtg cacttgtttg tcgggtcatc   4980

tttttcatgct ttttttttgtc ttggttgtga tgatgtggtc tggttgggcg gtcgttctag   5040

atcggagtag aattctgttt caaactacct ggtggattta ttaattttgg atctgtatgt   5100

gtgtgccata catattcata gttacgaatt gaagatgatg gatggaaata tcgatctagg   5160

ataggtatac atgttgatgc gggtttttact gatgcatata cagagatgct ttttgttcgc   5220

ttggttgtga tgatgtggtg tggttgggcg gtcgttcatt cgttctagat cggagtagaa   5280

tactgtttca aactacctgg tgtatttatt aattttggaa ctgtatgtgt gtgtcataca   5340

tcttcatagt tacgagttta agatggatgg aaatatcgat ctaggatagg tatacatgtt   5400

gatgtgggtt ttactgatgc atatacatga tggcatatgc agcatctatt catatgctct   5460

aaccttgagt acctatctat tataataaac aagtatgttt tataattatt ttgatcttga   5520

tatacttgga tgatggcata tgcagcagct atatgtggat ttttttagcc ctgccttcat   5580

acgctattta tttgcttggt actgtttctt ttgtcgatgc tcaccctgtt gtttggtgtt   5640

acttctgcag gtcgacttta acttagccta ggatccacac gacaccatgt cccccgagcg   5700

ccgccccgtc gagatccgcc cggccaccgc cgccgacatg gccgccgtgt gcgacatcgt   5760

gaaccactac atcgagacct ccaccgtgaa cttccgcacc gagccgcaga ccccgcagga   5820

gtggatcgac gacctggagc gcctccagga ccgctacccg tggctcgtgg ccgaggtgga   5880

gggcgtggtg gccggcatcg cctacgccgg cccgtggaag gcccgcaacg cctacgactg   5940

gaccgtggag tccaccgtgt acgtgtccca ccgccaccag cgcctcggcc tcggctccac   6000

cctctacacc cacctcctca agagcatgga ggcccagggc ttcaagtccg tggtggccgt   6060

gatcggcctc ccgaacgacc cgtccgtgcg cctccacgag gccctcggct acaccgcccg   6120

cggcaccctc cgcgccgccg gctacaagca cggcggctgg cacgacgtcg gcttctggca   6180

gcgcgacttc gagctgccgg ccccgccgcg cccggtgcgc ccggtgacgc agatctgagt   6240

cgaaacctag acttgtccat cttctggatt ggccaactta attaatgtat gaaataaaag   6300

gatgcacaca tagtgacatg ctaatcacta taatgtgggc atcaaagttg tgtgttatgt   6360

gtaattacta gttatctgaa taaaagagaa agagatcatc catatttctt atcctaaatg   6420

aatgtcacgt gtctttataa ttctttgatg aaccagatgc atttcattaa ccaaatccat   6480

atacatataa atattaatca tatataatta atatcaattg ggttagcaaa acaaatctag   6540

tctaggtgtg ttttgcgaat tgcggccgcc accgcggtgg agctcgaatt cattccgatt   6600

aatcgtggcc tcttgctctt caggatgaag agctatgttt aaacgtgcaa gcgctactag   6660

acaattcagt acattaaaaa cgtccgcaat gtgttattaa gttgtctaag cgtcaatttg   6720

tttacaccac aatatatcct gccaccagcc agccaacagc tccccgaccg gcagctcggc   6780
```

-continued

```
acaaaatcac cactcgatac aggcagccca tcagtccggg acggcgtcag cgggagagcc    6840

gttgtaaggc ggcagacttt gctcatgtta ccgatgctat tcggaagaac ggcaactaag    6900

ctgccgggtt tgaaacacgg atgatctcgc ggagggtagc atgttgattg taacgatgac    6960

agagcgttgc tgcctgtgat caaatatcat ctccctcgca gagatccgaa ttatcagcct    7020

tcttattcat ttctcgctta accgtgacag gctgtcgatc ttgagaacta tgccgacata    7080

ataggaaatc gctggataaa gccgctgagg aagctgagtg gcgctatttc tttagaagtg    7140

aacgttgacg atcgtcgacc gtaccccgat gaattaattc ggacgtacgt tctgaacaca    7200

gctggatact tacttgggcg attgtcatac atgacatcaa caatgtaccc gtttgtgtaa    7260

ccgtctcttg gaggttcgta tgacactagt ggttcccctc agcttgcgac tagatgttga    7320

ggcctaacat tttattagag agcaggctag ttgcttagat acatgatctt caggccgtta    7380

tctgtcaggg caagcgaaaa ttggccattt atgacgacca atgccccgca gaagctccca    7440

tctttgccgc catagacgcc gcgcccccct tttggggtgt agaacatcct tttgccagat    7500

gtggaaaaga agttcgttgt cccattgttg gcaatgacgt agtagccggc gaaagtgcga    7560

gacccatttg cgctatatat aagcctacga tttccgttgc gactattgtc gtaattggat    7620

gaactattat cgtagttgct ctcagagttg tcgtaatttg atggactatt gtcgtaattg    7680

cttatggagt tgtcgtagtt gcttggagaa atgtcgtagt tggatgggga gtagtcatag    7740

ggaagacgag cttcatccac taaaacaatt ggcaggtcag caagtgcctg ccccgatgcc    7800

atcgcaagta cgaggcttag aaccaccttc aacagatcgc gcatagtctt ccccagctct    7860

ctaacgcttg agttaagccg cgccgcgaag cggcgtcggc ttgaacgaat tgttagacat    7920

tatttgccga ctaccttggt gatctcgcct ttcacgtagt gaacaaattc ttccaactga    7980

tctgcgcgcg aggccaagcg atcttcttgt ccaagataag cctgcctagc ttcaagtatg    8040

acgggctgat actgggccgg caggcgctcc attgcccagt cggcagcgac atccttcggc    8100

gcgattttgc cggttactgc gctgtaccaa atgcgggaca acgtaagcac tacatttcgc    8160

tcatcgccag cccagtcggg cggcgagttc catagcgtta aggtttcatt tagcgcctca    8220

aatagatcct gttcaggaac cggatcaaag agttcctccg ccgctggacc taccaaggca    8280

acgctatgtt ctcttgcttt tgtcagcaag atagccagat caatgtcgat cgtggctggc    8340

tcgaagatac ctgcaagaat gtcattgcgc tgccattctc caaattgcag ttcgcgctta    8400

gctggataac gccacggaat gatgtcgtcg tgcacaacaa tggtgacttc tacagcgcgg    8460

agaatctcgc tctctccagg ggaagccgaa gtttccaaaa ggtcgttgat caaagctcgc    8520

cgcgttgttt catcaagcct tacagtcacc gtaaccagca aatcaatatc actgtgtggc    8580

ttcaggccgc catccactgc ggagccgtac aaatgtacgg ccagcaacgt cggttcgaga    8640

tggcgctcga tgacgccaac tacctctgat agttgagtcg atacttcggc gatcaccgct    8700

tccctcatga tgtttaactc ctgaattaag ccgcgccgcg aagcggtgtc ggcttgaatg    8760

aattgttagg cgtcatcctg tgctcccgag aaccagtacc agtacatcgc tgtttcgttc    8820

gagacttgag gtctagtttt atacgtgaac aggtcaatgc cgccgagagt aaagccacat    8880

tttgcgtaca aattgcaggc aggtacattg ttcgtttgtg tctctaatcg tatgccaagg    8940

agctgtctgc ttagtgccca cttttttcgca aattcgatga gactgtgcgc gactcctttg    9000

cctcggtgcg tgtgcgacac aacaatgtgt tcgatagagg ctagatcgtt ccatgttgag    9060
```

-continued

```
ttgagttcaa tcttcccgac aagctcttgg tcgatgaatg cgccatagca agcagagtct      9120

tcatcagagt catcatccga gatgtaatcc ttccggtagg ggctcacact tctggtagat      9180

agttcaaagc cttggtcgga taggtgcaca tcgaacactt cacgaacaat gaaatggttc      9240

tcagcatcca atgtttccgc cacctgctca gggatcaccg aaatcttcat atgacgccta      9300

acgcctggca cagcggatcg caaacctggc gcggcttttg cacaaaagg cgtgacaggt      9360

ttgcgaatcc gttgctgcca cttgttaacc cttttgccag atttggtaac tataatttat      9420

gttagaggcg aagtcttggg taaaaactgg cctaaaattg ctggggattt caggaaagta      9480

aacatcacct tccggctcga tgtctattgt agatatatgt agtgtatcta cttgatcggg      9540

ggatctgctg cctcgcgcgt ttcggtgatg acggtgaaaa cctctgacac atgcagctcc      9600

cggagacggt cacagcttgt ctgtaagcgg atgccgggag cagacaagcc cgtcagggcg      9660

cgtcagcggg tgttggcggg tgtcggggcg cagccatgac ccagtcacgt agcgatagcg      9720

gagtgtatac tggcttaact atgcggcatc agagcagatt gtactgagag tgcaccatat      9780

gcggtgtgaa ataccgcaca gatgcgtaag gagaaaatac cgcatcaggc gctcttccgc      9840

ttcctcgctc actgactcgc tgcgctcggt cgttcggctg cggcgagcgg tatcagctca      9900

ctcaaaggcg gtaatacggt tatccacaga atcaggggat aacgcaggaa agaacatgtg      9960

agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca     10020

taggctccgc cccCCtgacg agcatcacaa aaatcgacgc tcaagtcaga ggtggcgaaa     10080

cccgacagga ctataaagat accaggcgtt ccccctgga agctccctcg tgcgctctcc     10140

tgttccgacc ctgccgctta ccggatacct gtccgccttt ctcccttcgg gaagcgtggc     10200

gctttctcat agctcacgct gtaggtatct cagttcggtg taggtcgttc gctccaagct     10260

gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc gccttatccg gtaactatcg     10320

tcttgagtcc aacccggtaa gacacgactt atcgccactg gcagcagcca ctggtaacag     10380

gattagcaga gcgaggtatg taggcggtgc tacagagttc ttgaagtggt ggcctaacta     10440

cggctacact agaaggacag tatttggtat ctgcgctctg ctgaagccag ttaccttcgg     10500

aaaaagagtt ggtagctctt gatccggcaa acaaaccacc gctggtagcg gtggtttttt     10560

tgtttgcaag cagcagatta cgcgcagaaa aaaaggatct caagaagatc ctttgatctt     10620

ttctacgggg tctgacgctc agtggaacga aaactcacgt taagggattt tggtcatgag     10680

attatcaaaa aggatcttca cctagatcct tttaaattaa aaatgaagtt ttaaatcaat     10740

ctaaagtata tatgagtaaa cttggtctga cagttaccaa tgcttaatca gtgaggcacc     10800

tatctcagcg atctgtctat ttcgttcatc catagttgcc tgactccccg tcgtgtagat     10860

aactacgata cgggagggct taccatctgg ccccagtgct gcaatgatac cgcgagaccc     10920

acgctcaccg gctccagatt tatcagcaat aaaccagcca gccggaaggg ccgagcgcag     10980

aagtggtcct gcaactttat ccgcctccat ccagtctatt aattgttgcc gggaagctag     11040

agtaagtagt tcgccagtta atagtttgcg caacgttgtt gccattgctg caggggggggg     11100

gggggggggg gacttccatt gttcattcca cggacaaaaa cagagaaagg aaacgacaga     11160

ggccaaaaag cctcgctttc agcacctgtc gtttcctttc ttttcagagg gtattttaaa     11220

taaaaacatt aagttatgac gaagaagaac ggaaacgcct aaaccggaa aattttcata     11280

aatagcgaaa acccgcgagg tcgccgcccc gtaacctgtc ggatcaccgg aaaggacccg     11340
```

-continued

```
taaagtgata atgattatca tctacatatc acaacgtgcg tggaggccat caaaccacgt    11400

caaataatca attatgacgc aggtatcgta ttaattgatc tgcatcaact taacgtaaaa    11460

acaacttcag acaatacaaa tcagcgacac tgaatacggg gcaacctcat gtccccccccc    11520

cccccccccc tgcaggcatc gtggtgtcac gctcgtcgtt tggtatggct tcattcagct    11580

ccggttccca acgatcaagg cgagttacat gatcccccat gttgtgcaaa aaagcggtta    11640

gctccttcgg tcctccgatc gttgtcagaa gtaagttggc cgcagtgtta tcactcatgg    11700

ttatggcagc actgcataat tctcttactg tcatgccatc cgtaagatgc ttttctgtga    11760

ctggtgagta ctcaaccaag tcattctgag aatagtgtat gcggcgaccg agttgctctt    11820

gcccggcgtc aacacgggat aataccgcgc cacatagcag aactttaaaa gtgctcatca    11880

ttggaaaacg ttcttcgggg cgaaaactct caaggatctt accgctgttg agatccagtt    11940

cgatgtaacc cactcgtgca cccaactgat cttcagcatc ttttactttc accagcgttt    12000

ctgggtgagc aaaaacagga aggcaaaatg ccgcaaaaaa gggaataagg cgacacgga    12060

aatgttgaat actcatactc ttcctttttc aatattattg aagcatttat cagggttatt    12120

gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaaata ggggttccgc    12180

gcacatttcc ccgaaaagtg ccacctgacg tctaagaaac cattattatc atgacattaa    12240

cctataaaaa taggcgtatc acgaggccct ttcgtcttca agaattggtc gacgatcttg    12300

ctgcgttcgg atattttcgt ggagttcccg ccacagaccc ggattgaagg cgagatccag    12360

caactcgcgc cagatcatcc tgtgacggaa ctttggcgcg tgatgactgg ccaggacgtc    12420

ggccgaaaga gcgacaagca gatcacgctt ttcgacagcg tcggatttgc gatcgaggat    12480

ttttcggcgc tgcgctacgt ccgcgaccgc gttgagggat caagccacag cagcccactc    12540

gaccttctag ccgacccaga cgagccaagg gatctttttg gaatgctgct ccgtcgtcag    12600

gctttccgac gtttgggtgg ttgaacagaa gtcattatcg tacggaatgc caagcactcc    12660

cgaggggaac cctgtggttg gcatgcacat acaaatggac gaacggataa acctttttcac    12720

gcccttttaa atatccgtta ttctaataaa cgctctttttc tcttaggttt acccgccaat    12780

atatcctgtc aaacactgat agtttaaact gaaggcggga aacgacaatc tgatcatgag    12840

cggagaatta agggagtcac gttatgaccc ccgccgatga cgcgggacaa gccgttttac    12900

gtttggaact gacagaaccg caacgttgaa ggagccactc agcaagctgg tacgattgta    12960

atacgactca ctatagggcg aattgagcgc tgtttaaacg ctcttcaact ggaagagcg    13019
```

<210> SEQ ID NO 43
<211> LENGTH: 49765
<212> TYPE: DNA
<213> ORGANISM: artificial
<220> FEATURE:
<223> OTHER INFORMATION: vector

<400> SEQUENCE: 43

```
gggggggggg ggggggggtt ccattgttca ttccacggac aaaaacagag aaaggaaacg      60

acagaggcca aaaagctcgc tttcagcacc tgtcgtttcc tttctttttca gagggtattt     120

taaataaaaa cattaagtta tgacgaagaa gaacggaaac gccttaaacc ggaaaatttt     180

cataaatagc gaaaacccgc gaggtcgccg ccccgtaacc tgtcggatca ccggaaagga     240

cccgtaaagt gataatgatt atcatctaca tatcacaacg tgcgtggagg ccatcaaacc     300
```

-continued

```
acgtcaaata atcaattatg acgcaggtat cgtattaatt gatctgcatc aacttaacgt    360

aaaaacaact tcagacaata caaatcagcg acactgaata cggggcaacc tcatgtcccc    420

ccccccccc ccctgcagg catcgtggtg tcacgctcgt cgtttggtat ggcttcattc     480

agctccggtt cccaacgatc aaggcgagtt acatgatccc ccatgttgtg caaaaaagcg    540

gttagctcct tcggtcctcc gatcgttgtc agaagtaagt tggccgcagt gttatcactc    600

atggttatgg cagcactgca taattctctt actgtcatgc catccgtaag atgcttttct    660

gtgactggtg agtactcaac caagtcattc tgagaatagt gtatgcggcg accgagttgc    720

tcttgcccgg cgtcaacacg ggataatacc gcgccacata gcagaacttt aaaagtgctc    780

atcattggaa aacgttcttc ggggcgaaaa ctctcaagga tcttaccgct gttgagatcc    840

agttcgatgt aacccactcg tgcacccaac tgatcttcag catctttttac tttcaccagc    900

gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa aaaagggaat aagggcgaca    960

cggaaatgtt gaatactcat actcttcctt tttcaatatt attgaagcat ttatcagggt    1020

tattgtctca tgagcggata catatttgaa tgtatttaga aaaataaaca aataggggtt    1080

ccgcgcacat ttccccgaaa agtgccacct gacgtctaag aaaccattat tatcatgaca    1140

ttaacctata aaaataggcg tatcacgagg ccctttcgtc ttcaagaatt cggagctttt    1200

gccattctca ccggattcag tcgtcactca tggtgatttc tcacttgata accttatttt    1260

tgacgagggg aaattaatag gttgtattga tgttggacga gtcggaatcg cagaccgata    1320

ccaggatctt gccatcctat ggaactgcct cggtgagttt tctccttcat tacagaaacg    1380

gctttttcaa aaatatggta ttgataatcc tgatatgaat aaattgcagt ttcatttgat    1440

gctcgatgag tttttctaat cagaattggt taattggttg taacactggc agagcattac    1500

gctgacttga cgggacggcg ctttgttga ataaatcgaa cttttgctga gttgaaggat     1560

cagatcacgc atcttcccga caacgcagac cgttccgtgg caaagcaaaa gttcaaaatc    1620

accaactggt ccacctacaa caaagctctc atcaaccgtg ctccctcac tttctggctg     1680

gatgatgggg cgattcaggc ctggtatgag tcagcaacac cttcttcacg aggcagacct    1740

cagcgccaga aggccgccag agaggccgag cgcggccgtg aggcttggac gctagggcag    1800

ggcatgaaaa agcccgtagc gggctgctac gggcgtctga cgcggtggaa agggggaggg    1860

gatgttgtct acatggctct gctgtagtga gtgggttgcg ctccggcagc ggtcctgatc    1920

aatcgtcacc ctttctcggt ccttcaacgt tcctgacaac gagcctcctt ttcgccaatc    1980

catcgacaat caccgcgagt ccctgctcga acgctgcgtc cggaccggct tcgtcgaagg    2040

cgtctatcgc ggcccgcaac agcggcgaga gcggagcctg ttcaacggtg ccgccgcgct    2100

cgccggcatc gctgtcgccg gctgctcct caagcacggc cccaacagtg aagtagctga     2160

ttgtcatcag cgcattgacg gcgtccccgg ccgaaaaacc gcctcgcag aggaagcgaa     2220

gctgcgcgtc ggccgtttcc atctgcggtg cgccggtcg cgtgccggca tggatgcgcg     2280

cgccatcgcg gtaggcgagc agcgcctgcc tgaagctgcg ggcattcccg atcagaaatg    2340

agcgccagtc gtcgtcggct ctcggcaccg aatgcgtatg attctccgcc agcatggctt    2400

cggccagtgc gtcgagcagc gcccgcttgt tcctgaagtg ccagtaaagc gccggctgct    2460

gaacccccaa ccgttccgcc agtttgcgtg tcgtcagacc gtctacgccg acctcgttca    2520

acaggtccag ggcggcacgg atcactgtat tcggctgcaa ctttgtcatg cttgacactt    2580
```

-continued

```
tatcactgat aaacataata tgtccaccaa cttatcagtg ataaagaatc cgcgcgttca    2640

atcggaccag cggaggctgg tccggaggcc agacgtgaaa cccaacatac ccctgatcgt    2700

aattctgagc actgtcgcgc tcgacgctgt cggcatcggc ctgattatgc cggtgctgcc    2760

gggcctcctg cgcgatctgg ttcactcgaa cgacgtcacc gcccactatg gcattctgct    2820

ggcgctgtat gcgttggtgc aatttgcctg cgcacctgtg ctgggcgcgc tgtcggatcg    2880

tttcgggcgg cggccaatct tgctcgtctc gctggccggc gccactgtcg actacgccat    2940

catggcgaca gcgcctttcc tttgggttct ctatatcggg cggatcgtgg ccggcatcac    3000

cggggcgact gggcggtag ccggcgctta tattgccgat atcactgatg gcgatgagcg    3060

cgcgcggcac ttcggcttca tgagcgcctg tttcgggttc gggatggtcg cgggacctgt    3120

gctcggtggg ctgatgggcg gtttctcccc ccacgctccg ttcttcgccg cggcagcctt    3180

gaacggcctc aatttcctga cgggctgttt ccttttgccg gagtcgcaca aaggcgaacg    3240

ccggccgtta cgccgggagg ctctcaaccc gctcgcttcg ttccggtggg cccgggggcat   3300

gaccgtcgtc gccgccctga tggcggtctt cttcatcatg caacttgtcg gacaggtgcc    3360

ggccgcgctt tgggtcattt tcggcgagga tcgctttcac tgggacgcga ccacgatcgg    3420

catttcgctt gccgcatttg gcattctgca ttcactcgcc caggcaatga tcaccggccc    3480

tgtagccgcc cggctcggcg aaaggcgggc actcatgctc ggaatgattg ccgacggcac    3540

aggctacatc ctgcttgcct tcgcgacacg gggatggatg gcgttcccga tcatggtcct    3600

gcttgcttcg ggtggcatcg gaatgccggc gctgcaagca atgttgtcca ggcaggtgga    3660

tgaggaacgt caggggcagc tgcaaggctc actggcggcg ctcaccagcc tgacctcgat    3720

cgtcggaccc ctcctcttca cggcgatcta tgcggcttct ataacaacgt ggaacgggtg    3780

ggcatggatt gcaggcgctg ccctctactt gctctgcctg ccggcgctgc gtcgcgggct    3840

ttggagcggc gcaggcaac gagccgatcg ctgatcgtgg aaacgatagg cctatgccat    3900

gcgggtcaag gcgacttccg gcaagctata cgcgccctag gagtgcggtt ggaacgttgg    3960

cccagccaga tactcccgat cacgagcagg acgccgatga tttgaagcgc actcagcgtc    4020

tgatccaaga acaaccatcc tagcaacacg gcggtccccg ggctgagaaa gcccagtaag    4080

gaaacaactg taggttcgag tcgcgagatc ccccggaacc aaaggaagta ggttaaaccc    4140

gctccgatca ggccgagcca cgccaggccg agaacattgg ttcctgtagg catcgggatt    4200

ggcggatcaa acactaaagc tactggaacg agcagaagtc ctccggccgc cagttgccag    4260

gcggtaaagg tgagcagagg cacggaggt tgccacttgc gggtcagcac ggttccgaac    4320

gccatggaaa ccgcccccgc caggcccgct gcgacgccga caggatctag cgctgcgttt    4380

ggtgtcaaca ccaacagcgc cacgcccgca gttccgcaaa tagcccccag gaccgccatc    4440

aatcgtatcg ggctacctag cagagcggca gagatgaaca cgaccatcag cggctgcaca    4500

gcgcctaccg tcgccgcgac cccgccggcc aggcggtaga ccgaaataaa caacaagctc    4560

cagaatagcg aaatattaag tgcgccgagg atgaagatgc gcatccacca gattcccgtt    4620

ggaatctgtc ggacgatcat cacgagcaat aaacccgccg gcaacgcccg cagcagcata    4680

ccggcgaccc ctcggcctcg ctgttcgggc tccacgaaaa cgccggacag atgcgccttg    4740

tgagcgtcct tggggccgtc ctcctgtttg aagaccgaca gcccaatgat ctcgccgtcg    4800

atgtaggcgc cgaatgccac ggcatctcgc aaccgttcag cgaacgcctc catgggcttt    4860
```

-continued

```
ttctcctcgt gctcgtaaac ggacccgaac atctctggag ctttcttcag ggccgacaat      4920

cggatctcgc ggaaatcctg cacgtcggcc gctccaagcc gtcgaatctg agccttaatc      4980

acaattgtca attttaatcc tctgtttatc ggcagttcgt agagcgcgcc gtgcgtcccg      5040

agcgatactg agcgaagcaa gtgcgtcgag cagtgcccgc ttgttcctga aatgccagta      5100

aagcgctggc tgctgaaccc ccagccggaa ctgaccccac aaggccctag cgtttgcaat      5160

gcaccaggtc atcattgacc caggcgtgtt ccaccaggcc gctgcctcgc aactcttcgc      5220

aggcttcgcc gacctgctcg cgccacttct tcacgcgggt ggaatccgat ccgcacatga      5280

ggcggaaggt ttccagcttg agcgggtacg gctcccggtg cgagctgaaa tagtcgaaca      5340

tccgtcgggc cgtcggcgac agcttgcggt acttctccca tatgaatttc gtgtagtggt      5400

cgccagcaaa cagcacgacg atttcctcgt cgatcaggac ctggcaacgg acgttttct      5460

tgccacggtc caggacgcgg aagcggtgca gcagcgacac cgattccagg tgcccaacgc      5520

ggtcggacgt gaagcccatc gccgtcgcct gtaggcgcga caggcattcc tcggccttcg      5580

tgtaataccg gccattgatc gaccagccca ggtcctggca aagctcgtag aacgtgaagg      5640

tgatcggctc gccgataggg gtgcgcttcg cgtactccaa cacctgctgc cacaccagtt      5700

cgtcatcgtc ggcccgcagc tcgacgccgg tgtaggtgat cttcacgtcc ttgttgacgt      5760

ggaaaatgac cttgtttgc agcgcctcgc gcgggatttt cttgttgcgc gtggtgaaca      5820

gggcagagcg ggccgtgtcg tttggcatcg ctcgcatcgt gtccggccac ggcgcaatat      5880

cgaacaagga aagctgcatt ccttgatct gctgcttcgt gtgtttcagc aacgcggcct      5940

gcttggcctc gctgacctgt tttgccaggt cctcgccggc ggttttttcgc ttcttggtcg      6000

tcatagttcc tcgcgtgtcg atggtcatcg acttcgccaa acctgccgcc tcctgttcga      6060

gacgacgcga acgctccacg gcggccgatg gcgcgggcag ggcagggga gccagttgca      6120

cgctgtcgcg ctcgatcttg gccgtagctt gctggaccat cgagccgacg gactggaagg      6180

tttcgcgggg cgcacgcatg acggtgcggc ttgcgatggt ttcggcatcc tcggcggaaa      6240

accccgcgtc gatcagttct tgcctgtatg ccttccggtc aaacgtccga ttcattcacc      6300

ctccttgcgg gattgccccg actcacgccg gggcaatgtg cccttattcc tgatttgacc      6360

cgcctggtgc cttggtgtcc agataatcca ccttatcggc aatgaagtcg gtcccgtaga      6420

ccgtctggcc gtccttctcg tacttggtat tccgaatctt gccctgcacg aataccagcg      6480

acccccttgcc caaatacttg ccgtgggcct cggcctgaga gccaaaacac ttgatgcgga      6540

agaagtcggt gcgctcctgc ttgtcgccgg catcgttgcg ccactcttca ttaaccgcta      6600

tatcgaaaat tgcttgcggc ttgttagaat tgccatgacg tacctcggtg tcacgggtaa      6660

gattaccgat aaactggaac tgattatggc tcatatcgaa agtctccttg agaaaggaga      6720

ctctagtttta gctaaacatt ggttccgctg tcaagaactt tagcggctaa aattttgcgg      6780

gccgcgacca aagtgcgag gggcggcttc cgctgtgtac aaccagatat ttttcaccaa      6840

catccttcgt ctgctcgatg agcggggcat gacgaaacat gagctgtcgg agagggcagg      6900

ggtttcaatt tcgttttat cagacttaac caacggtaag gccaacccct cgttgaaggt      6960

gatggaggcc attgccgacg ccctggaaac tccctacct cttctcctgg agtccaccga      7020

ccttgaccgc gaggcactcg cggagattgc gggtcatcct ttcaagagca gcgtgccgcc      7080

cggatacgaa cgcatcagtg tggttttgcc gtcacataag gcgtttatcg taaagaaatg      7140
```

-continued

```
gggcgacgac acccgaaaaa agctgcgtgg aaggctctga cgccaagggt tagggcttgc    7200

acttccttct ttagccgcta aaacggcccc ttctctgcgg gccgtcggct cgcgcatcat    7260

atcgacatcc tcaacggaag ccgtgccgcg aatggcatcg ggcgggtgcg ctttgacagt    7320

tgttttctat cagaacccct acgtcgtgcg gttcgattag ctgtttgtct tgcaggctaa    7380

acactttcgg tatatcgttt gcctgtgcga taatgttgct aatgatttgt tgcgtagggg    7440

ttactgaaaa gtgagcggga aagaagagtt tcagaccatc aaggagcggg ccaagcgcaa    7500

gctggaacgc gacatgggtg cggacctgtt ggccgcgctc aacgacccga aaaccgttga    7560

agtcatgctc aacgcggacg gcaaggtgtg gcacgaacgc cttggcgagc cgatgcggta    7620

catctgcgac atgcggccca gccagtcgca ggcgattata gaaacggtgg ccggattcca    7680

cggcaaagag gtcacgcggc attcgcccat cctggaaggc gagttcccct tggatggcag    7740

ccgctttgcc ggccaattgc cgccggtcgt ggccgcgcca acctttgcga tccgcaagcg    7800

cgcggtcgcc atcttcacgc tggaacagta cgtcgaggcg ggcatcatga cccgcgagca    7860

atacgaggtc attaaaagcg ccgtcgcggc gcatcgaaac atcctcgtca ttggcggtac    7920

tggctcgggc aagaccacgc tcgtcaacgg gatcatcaat gaaatggtcg ccttcaaccc    7980

gtctgagcgc gtcgtcatca tcgaggacac cggcgaaatc cagtgcgccg cagagaacgc    8040

cgtccaatac cacaccagca tcgacgtctc gatgacgctg ctgctcaaga caacgctgcg    8100

tatgcgcccc gaccgcatcc tggtcggtga ggtacgtggc cccgaagccc ttgatctgtt    8160

gatggcctgg aacaccgggc atgaaggagg tgccgccacc ctgcacgcaa acaaccccaa    8220

agcgggcctg agccggctcg ccatgcttat cagcatgcac ccggattcac cgaaacccat    8280

tgagccgctg attggcgagg cggttcatgt ggtcgtccat atcgccagga cccctagcgg    8340

ccgtcgagtg caagaaattc tcgaagttct tggttacgag aacggccagt acatcaccaa    8400

aaccctgtaa ggagtatttc caatgacaac ggctgttccg ttccgtctga ccatgaatcg    8460

cggcattttg ttctaccttg ccgtgttctt cgttctcgct ctcgcgttat ccgcgcatcc    8520

ggcgatggcc tcggaaggca ccggcggcag cttgccatat gagagctggc tgacgaacct    8580

gcgcaactcc gtaaccggcc cggtggcctt cgcgctgtcc atcatcggca tcgtcgtcgc    8640

cggcggcgtg ctgatcttcg gcggcgaact caacgccttc ttccgaaccc tgatcttcct    8700

ggttctggtg atggcgctgc tggtcggcgc gcagaacgtg atgagcacct tcttcggtcg    8760

tggtgccgaa atcgcggccc tcggcaacgg ggcgctgcac caggtgcaag tcgcggcggc    8820

ggatgccgtg cgtgcggtag cggctggacg gctcgcctaa tcatggctct cgcacgatc    8880

cccatccgtc gcgcaggcaa ccagaaaaac ctgttcatgg gtggtgatcg tgaactggtg    8940

atgttctcgg gcctgatggc gtttgcgctg attttcagcg cccaagagct gcgggccacc    9000

gtggtcggtc tgatcctgtg gttcggggcg ctctatgcgt tccgaatcat ggcgaaggcc    9060

gatccgaaga tgcggttcgt gtacctgcgc caccgccggt acaagccgta ttacccggcc    9120

cgctcgaccc cgttccgcga gaacaccaat agccaaggga agcaataccg atgatccaag    9180

caattgcgat tgcaatcgcg ggcctcggcc cgcttctgtt gttcatcctc tttgcccgca    9240

tccgcgcggt cgatgccgaa ctgaaactga aaaagcatcg ttccaaggac gccggcctgg    9300

ccgatctgct caactacgcc gctgtcgtcg atgacggcgt aatcgtgggc aagaacggca    9360

gctttatggc tgcctggctg tacaagggcg atgacaacgc aagcagcacc gaccagcagc    9420
```

-continued

```
gcgaagtagt gtccgcccgc atcaaccagg ccctcgcggg cctgggaagt gggtggatga    9480

tccatgtgga cgccgtgcgg cgtcctgctc cgaactacgc ggagcggggc ctgtcggcgt    9540

tccctgaccg tctgacggca gcgattgaag aagagcgctc ggtcttgcct tgctcgtcgg    9600

tgatgtactt caccagctcc gcgaagtcgc tcttcttgat ggagcgcatg gggacgtgct    9660

tggcaatcac gcgcaccccc cggccgtttt agcggctaaa aaagtcatgg ctctgccctc    9720

gggcggacca cgcccatcat gaccttgcca agctcgtcct gcttctcttc gatcttcgcc    9780

agcagggcga ggatcgtggc atcaccgaac cgcgccgtgc gcgggtcgtc ggtgagccag    9840

agtttcagca ggccgcccag gcggcccagg tcgccattga tgcgggccag ctcgcggacg    9900

tgctcatagt ccacgacgcc cgtgattttg tagccctggc cgacggccag caggtaggcc    9960

gacaggctca tgccggccgc cgccgccttt tcctcaatcg ctcttcgttc gtctggaagg   10020

cagtacacct tgataggtgg gctgcccttc ctggttggct tggtttcatc agccatccgc   10080

ttgccctcat ctgttacgcc ggcggtagcc ggccagcctc gcagagcagg attcccgttg   10140

agcaccgcca ggtgcgaata agggacagtg aagaaggaac acccgctcgc gggtgggcct   10200

acttcaccta tcctgcccgg ctgacgccgt tggatacacc aaggaaagtc tacacgaacc   10260

ctttggcaaa atcctgtata tcgtgcgaaa aaggatggat ataccgaaaa aatcgctata   10320

atgaccccga agcagggtta tgcagcggaa aagcgctgct ccctgctgt tttgtggaat     10380

atctaccgac tggaaacagg caaatgcagg aaaattactga actgagggga caggcgagag   10440

acgatgccaa agagctacac cgacgagctg gccgagtggg ttgaatcccg cgcggccaag   10500

aagcgccggc gtgatgaggc tgcggttgcg ttcctggcgg tgagggcgga tgtcgaggcg   10560

gcgttagcgt ccggctatgc gctcgtcacc atttgggagc acatgcggga aacggggaag   10620

gtcaagttct cctacgagac gttccgctcg cacgccaggc ggcacatcaa ggccaagccc   10680

gccgatgtgc ccgcaccgca ggccaaggct gcggaacccg cgccggcacc caagacgccg   10740

gagccacggc ggccgaagca gggggggcaag gctgaaaagc cggcccccgc tgcggccccg   10800

accggcttca ccttcaaccc aacaccggac aaaaaggatc tactgtaatg gcgaaaattc   10860

acatggtttt gcagggcaag ggcggggtcg gcaagtcggc catcgccgcg atcattgcgc   10920

agtacaagat ggacaagggg cagacaccct tgtgcatcga caccgacccg gtgaacgcga   10980

cgttcgaggg ctacaaggcc ctgaacgtcc gccggctgaa catcatggcc ggcgacgaaa   11040

ttaactcgcg caacttcgac accctggtcg agctgattgc gccgaccaag gatgacgtgg   11100

tgatcgacaa cggtgccagc tcgttcgtgc ctctgtcgca ttacctcatc agcaaccagg   11160

tgccggctct gctgcaagaa atggggcatg agctggtcat ccataccgtc gtcaccggcg   11220

gccaggctct cctggacacg gtgagcggct cgcccagct cgccagccag ttcccggccg    11280

aagcgctttt cgtggtctgg ctgaacccgt attggggggcc tatcgagcat gagggcaaga   11340

gctttgagca gatgaaggcg tacacggcca acaaggcccg cgtgtcgtcc atcatccaga   11400

ttccggccct caaggaagaa acctacggcc gcgatttcag cgacatgctg caagagcggc   11460

tgacgttcga ccaggcgctg gccgatgaat cgctcacgat catgacgcgc caacgcctca   11520

agatcgtgcg gcgcggcctg tttgaacagc tcgacgcggc ggccgtgcta tgagcgacca   11580

gattgaagag ctgatccggg agattgcggc caagcacggc atcgccgtcg ccgcgacga    11640

cccggtgctg atcctgcata ccatcaacgc ccggctcatg gccgacagtg cggccaagca   11700
```

-continued

```
agaggaaatc cttgccgcgt tcaaggaaga gctggaaggg atcgcccatc gttgggcga    11760

ggacgccaag gccaaagcgg agcggatgct gaacgcggcc ctggcggcca gcaaggacgc    11820

aatggcgaag gtaatgaagg acagcgccgc gcaggcggcc gaagcgatcc gcagggaaat    11880

cgacgacggc cttggccgcc agctcgcggc caaggtcgcg gacgcgcggc gcgtggcgat    11940

gatgaacatg atcgccgcg gcatggtgtt gttcgcggcc gccctggtgg tgtgggcctc     12000

gttatgaatc gcagaggcgc agatgaaaaa gcccggcgtt gccgggcttt gtttttgcgt     12060

tagctgggct tgtttgacag gcccaagctc tgactgcgcc cgcgctcgcg ctcctgggcc     12120

tgtttcttct cctgctcctg cttgcgcatc agggcctggt gccgtcgggc tgcttcacgc     12180

atcgaatccc agtcgccggc cagctcggga tgctccgcgc gcatcttgcg cgtcgccagt     12240

tcctcgatct tgggcgcgtg aatgcccatg ccttccttga tttcgcgcac catgtccagc     12300

cgcgtgtgca gggtctgcaa gcgggcttgc tgttgggcct gctgctgctg ccaggcggcc     12360

tttgtacgcg gcagggacag caagccgggg gcattggact gtagctgctg caaacgcgcc     12420

tgctgacggt ctacgagctg ttctaggcgg tcctcgatgc gctccacctg gtcatgcttt     12480

gcctgcacgt agagcgcaag ggtctgctgg taggtctgct cgatgggcgc ggattctaag     12540

agggcctgct gttccgtctc ggcctcctgg gccgcctgta gcaaatcctc gccgctgttg     12600

ccgctggact gctttactgc cggggactgc tgttgccctg ctcgcgccgt cgtcgcagtt     12660

cggcttgccc ccactcgatt gactgcttca tttcgagccg cagcgatgcg atctcggatt     12720

gcgtcaacgg acggggcagc gcggaggtgt ccggcttctc cttgggtgag tcggtcgatg     12780

ccatagccaa aggtttcctt ccaaaatgcg tccattgctg gaccgtgttt ctcattgatg     12840

cccgcaagca tcttcggctt gaccgccagg tcaagcgcgc cttcatgggc ggtcatgacg     12900

gacgccgcca tgaccttgcc gccgttgttc tcgatgtagc cgcgtaatga ggcaatggtg     12960

ccgcccatcg tcagcgtgtc atcgacaacg atgtacttct ggccggggat cacctccccc     13020

tcgaaagtcg ggttgaacgc caggcgatga tctgaaccgg ctccggttcg ggcgaccttc     13080

tcccgctgca caatgtccgt ttcgacctca aggccaaggc ggtcggccag aacgaccgcc     13140

atcatggccg gaatcttgtt gttccccgcc gcctcgacgg cgaggactgg aacgatgcgg     13200

ggcttgtcgt cgccgatcag cgtcttgagc tgggcaacag tgtcgtccga aatcaggcgc     13260

tcgaccaaat taagcgccgc ttccgcgtcg ccctgcttcg cagcctggta ttcaggctcg     13320

ttggtcaaag aaccaaggtc gccgttgcga accaccttcg ggaagtctcc ccacggtgcg     13380

cgctcggctc tgctgtagct gctcaagacg cctccctttt tagccgctaa aactctaacg     13440

agtgcgcccg cgactcaact tgacgctttc ggcacttacc tgtgccttgc cacttgcgtc     13500

ataggtgatg cttttcgcac tcccgatttc aggtacttta tcgaaatctg accgggcgtg     13560

cattacaaag ttcttcccca cctgttggta aatgctgccg ctatctgcgt ggacgatgct     13620

gccgtcgtgg cgctgcgact tatcggcctt ttgggccata tagatgttgt aaatgccagg     13680

tttcagggcc ccggctttat ctaccttctg gttcgtccat gcgccttggt tctcggtctg     13740

gacaattctt tgcccattca tgaccaggag gcggtgtttc attgggtgac tcctgacggt     13800

tgcctctggt gttaaacgtg tcctggtcgc ttgccggcta aaaaaaagcc gacctcggca     13860

gttcgaggcc ggctttccct agagccgggc gcgtcaaggt tgttccatct attttagtga     13920

actgcgttcg atttatcagt tactttcctc ccgctttgtg tttcctccca ctcgtttccg     13980
```

-continued

```
cgtctagccg acccctcaac atagcggcct cttcttgggc tgcctttgcc tcttgccgcg   14040

cttcgtcacg ctcggcttgc accgtcgtaa agcgctcggc ctgcctggcc gcctcttgcg   14100

ccgccaactt cctttgctcc tggtgggcct cggcgtcggc ctgcgccttc gctttcaccg   14160

ctgccaactc cgtgcgcaaa ctctccgctt cgcgcctggt ggcgtcgcgc tcgccgcgaa   14220

gcgcctgcat ttcctggttg gccgcgtcca gggtcttgcg gctctcttct ttgaatgcgc   14280

gggcgtcctg gtgagcgtag tccagctcgg cgcgcagctc ctgcgctcga cgctccacct   14340

cgtcggcccg ctgcgtcgcc agcgcggccc gctgctcggc tcctgccagg gcggtgcgtg   14400

cttcggccag ggcttgccgc tggcgtgcgg ccagctcggc cgcctcggcg gcctgctgct   14460

ctagcaatgt aacgcgcgcc tgggcttctt ccagctcgcg ggcctgcgcc tcgaaggcgt   14520

cggccagctc cccgcgcacg gcttccaact cgttgcgctc acgatcccag ccggcttgcg   14580

ctgcctgcaa cgattcattg gcaagggcct gggcggcttg ccagagggcg gccacggcct   14640

ggttgccggc ctgctgcacc gcgtccggca cctggactgc cagcggggcg gcctgcgccg   14700

tgcgctggcg tcgccattcg cgcatgccgg cgctggcgtc gttcatgttg acgcgggcgg   14760

ccttacgcac tgcatccacg gtcgggaagt tctcccggtc gccttgctcg aacagctcgt   14820

ccgcagccgc aaaaatgcgg tcgcgcgtct cttttgttcag ttccatgttg gctccggtaa   14880

ttggtaagaa taataatact cttacctacc ttatcagcgc aagagtttag ctgaacagtt   14940

ctcgacttaa cggcaggttt tttagcggct gaagggcagg caaaaaaagc cccgcacggt   15000

cggcgggggc aaagggtcag cgggaagggg attagcgggc gtcgggcttc ttcatgcgtc   15060

ggggccgcgc ttcttgggat ggagcacgac gaagcgcgca cgcgcatcgt cctcggccct   15120

atcggcccgc gtcgcggtca ggaacttgtc gcgcgctagg tcctccctgg tgggcaccag   15180

gggcatgaac tcggcctgct cgatgtaggt ccactccatg accgcatcgc agtcgaggcc   15240

gcgttccttc accgtctctt gcaggtcgcg gtacgcccgc tcgttgagcg ctggtaacg   15300

ggccaattgg tcgtaaatgg ctgtcggcca tgagcggcct ttcctgttga gccagcagcc   15360

gacgacgaag ccggcaatgc aggcccctgg cacaaccagg ccgacgccgg gggcagggga   15420

tggcagcagc tcgccaacca ggaaccccgc cgcgatgatg ccgatgccgg tcaaccagcc   15480

cttgaaacta tccggccccg aaacaccccct gcgcattgcc tggatgctgc gccggatagc   15540

ttgcaacatc aggagccgtt tcttttgttc gtcagtcatg gtccgccctc accagttgtt   15600

cgtatcggtg tcggacgaac tgaaatcgca agagctgccg gtatcggtcc agccgctgtc   15660

cgtgtcgctg ctgccgaagc acggcgaggg gtccgcgaac gccgcagacg gcgtatccgg   15720

ccgcagcgca tcgcccagca tggcccccggt cagcgagccg ccggccaggt agcccagcat   15780

ggtgctgttg gtcgcccccgg ccaccagggc cgacgtgacg aaatcgccgt cattccctct   15840

ggattgttcg ctgctcggcg gggcagtgcg ccgcgccggc ggcgtcgtgg atggctcggg   15900

ttggctggcc tgcgacggcc ggcgaaaggt gcgcagcagc tcgttatcga ccggctgcgg   15960

cgtcgggggcc gccgccttgc gctgcggtcg gtgttccttc ttcggctcgc gcagcttgaa   16020

cagcatgatc gcgaaaacca gcagcaacgc cgcgcctacg cctcccgcga tgtagaacag   16080

catcggattc attcttcggt cctccttgta gcggaaccgt gtctgtgcg gcgcgggtgg   16140

cccgcgccgc tgtctttggg gatcagccct cgatgagcgc gaccagtttc acgtcggcaa   16200

ggttcgcctc gaactcctgg ccgtcgtcct cgtacttcaa ccaggcatag ccttccgccg   16260
```

-continued

```
gcggccgacg gttgaggata aggcgggcag ggcgctcgtc gtgctcgacc tggacgatgg    16320

ccttttttcag cttgtccggg tccggctcct tcgcgccctt ttccttggcg tccttaccgt    16380

cctggtcgcc gtcctcgccg tcctggccgt cgccggcctc cgcgtcacgc tcggcatcag    16440

tctggccgtt gaaggcatcg acggtgttgg gatcgcggcc cttctcgtcc aggaactcgc    16500

gcagcagctt gaccgtgccg cgcgtgattt cctgggtgtc gtcgtcaagc cacgcctcga    16560

cttcctccgg gcgcttcttg aaggccgtca ccagctcgtt caccacggtc acgtcgcgca    16620

cgcggccggt gttgaacgca tcggcgatct tctccggcag gtccagcagc gtgacgtgct    16680

gggtgatgaa cgccggcgac ttgccgattt ccttggcgat atcgcctttc ttcttgccct    16740

tcgccagctc gcggccaatg aagtcggcaa tttcgcgcgg ggtcagctcg ttgcgttgca    16800

ggttctcgat aacctggtcg gcttcgttgt agtcgttgtc gatgaacgcc gggatggact    16860

tcttgccggc ccacttcgag ccacggtagc ggcgggcgcc gtgattgatg atatagcggc    16920

ccggctgctc ctggttctcg cgcaccgaaa tgggtgactt caccccgcgc tctttgatcg    16980

tggcaccgat ttccgcgatg ctctccgggg aaaagccggg gttgtcggcc gtccgcggct    17040

gatgcggatc ttcgtcgatc aggtccaggt ccagctcgat agggccggaa ccgccctgag    17100

acgccgcagg agcgtccagg aggctcgaca ggtcgccgat gctatccaac cccaggccgg    17160

acggctgcgc cgcgcctgcg gcttcctgag cggccgcagc ggtgtttttc ttggtggtct    17220

tggcttgagc cgcagtcatt gggaaatctc catcttcgtg aacacgtaat cagccagggc    17280

gcgaacctct ttcgatgcct tgcgcgcggc cgtttttcttg atcttccaga ccggcacacc    17340

ggatgcgagg gcatcggcga tgctgctgcg caggccaacg gtggccggaa tcatcatctt    17400

ggggtacgcg gccagcagct cggcttggtg gcgcgcgtgg cgcggattcc gcgcatcgac    17460

cttgctgggc accatgccaa ggaattgcag cttggcgttc ttctggcgca cgttcgcaat    17520

ggtcgtgacc atcttcttga tgccctggat gctgtacgcc tcaagctcga tgggggacag    17580

cacatagtcg gccgcgaaga gggcggccgc caggccgacg ccaagggtcg gggccgtgtc    17640

gatcaggcac acgtcgaagc cttggttcgc cagggccttg atgttcgccc cgaacagctc    17700

gcgggcgtcg tccagcgaca gccgttcggc gttcgccagt accgggttgg actcgatgag    17760

ggcgaggcgc gcggcctggc cgtcgccggc tgcgggtgcg gtttcggtcc agccgccggc    17820

agggacagcg ccgaacagct tgcttgcatg caggccggta gcaaagtcct tgagcgtgta    17880

ggacgcattg ccctggggggt ccaggtcgat cacggcaacc cgcaagccgc gctcgaaaaa    17940

gtcgaaggca agatgcacaa gggtcgaagt cttgccgacg ccgcctttct ggttggccgt    18000

gaccaaagtt ttcatcgttt ggtttcctgt tttttcttgg cgtccgcttc ccacttccgg    18060

acgatgtacg cctgatgttc cggcagaacc gccgttaccc gcgcgtaccc ctcgggcaag    18120

ttcttgtcct cgaacgcggc ccacacgcga tgcaccgctt gcgacactgc gccctggtc    18180

agtcccagcg acgttgcgaa cgtcgcctgt ggcttccat cgactaagac gccccgcgct    18240

atctcgatgg tctgctgccc cacttccagc ccctggatcg cctcctggaa ctggctttcg    18300

gtaagccgtt tcttcatgga taacaccat aatttgctcc gcgccttggt tgaacatagc    18360

ggtgacagcc gccagcacat gagagaagtt tagctaaaca tttctcgcac gtcaacacct    18420

ttagccgcta aaactcgtcc ttggcgtaac aaaacaaaag cccggaaacc gggctttcgt    18480

ctcttgccgc ttatggctct gcacccggct ccatcaccaa caggtcgcgc acgcgcttca    18540
```

-continued

```
ctcggttgcg gatcgacact gccagcccaa caaagccggt tgccgccgcc gccaggatcg   18600

cgccgatgat gccggccaca ccggccatcg cccaccaggt cgccgccttc cggttccatt   18660

cctgctggta ctgcttcgca atgctggacc tcggctcacc ataggctgac cgctcgatgg   18720

cgtatgccgc ttctcccctt ggcgtaaaac ccagcgccgc aggcggcatt gccatgctgc   18780

ccgccgcttt cccgaccacg acgcgcgcac caggcttgcg gtccagacct tcggccacgg   18840

cgagctgcgc aaggacataa tcagccgccg acttggctcc acgcgcctcg atcagctctt   18900

gcactcgcgc gaaatccttg gcctccacgg ccgccatgaa tcgcgcacgc ggcgaaggct   18960

ccgcagggcc ggcgtcgtga tcgccgccga gaatgccctt caccaagttc gacgacacga   19020

aaatcatgct gacggctatc accatcatgc agacggatcg cacgaacccg ctgaattgaa   19080

cacgagcacg gcacccgcga ccactatgcc aagaatgccc aaggtaaaaa ttgccggccc   19140

cgccatgaag tccgtgaatg ccccgacggc cgaagtgaag ggcaggccgc cacccaggcc   19200

gccgccctca ctgcccggca cctggtcgct gaatgtcgat gccagcacct gcggcacgtc   19260

aatgcttccg ggcgtcgcgc tcgggctgat cgcccatccc gttactgccc cgatcccggc   19320

aatggcaagg actgccagcg ctgccatttt tggggtgagg ccgttcgcgg ccgaggggcg   19380

cagcccctgg ggggatggga ggcccgcgtt agcgggccgg gagggttcga gaaggggggg   19440

cacccccctt cggcgtgcgc ggtcacgcgc acagggcgca gccctggtta aaaacaaggt   19500

ttataaatat tggtttaaaa gcaggttaaa agacaggtta gcggtggccg aaaaacgggc   19560

ggaaacccctt gcaaatgctg gattttctgc ctgtggacag cccctcaaat gtcaataggt   19620

gcgcccctca tctgtcagca ctctgcccct caagtgtcaa ggatcgcgcc cctcatctgt   19680

cagtagtcgc gcccctcaag tgtcaatacc gcagggcact tatccccagg cttgtccaca   19740

tcatctgtgg aaactcgcg taaaatcagg cgttttcgcc gatttgcgag gctggccagc   19800

tccacgtcgc cggccgaaat cgagcctgcc cctcatctgt caacgccgcg ccgggtgagt   19860

cggcccctca agtgtcaacg tccgccccctc atctgtcagt gagggccaag ttttccgcga   19920

ggtatccaca acgccggcgg ccgcggtgtc tcgcacacgg cttcgacggc gtttctggcg   19980

cgtttgcagg gccatagacg gccgccagcc cagcggcgag ggcaaccagc ccggtgagcg   20040

tcggaaaggc gctggaagcc ccgtagcgac gcggagaggg gcgagacaag ccaaggcgc   20100

aggctcgatg cgcagcacga catagccggt tctcgcaagg acgagaattt ccctgcggtg   20160

cccctcaagt gtcaatgaaa gtttccaacg cgagccattc gcgagagcct tgagtccacg   20220

ctagatgaga gctttgttgt aggtggacca gttggtgatt ttgaactttt gctttgccac   20280

ggaacggtct gcgttgtcgg aagatgcgt gatctgatcc ttcaactcag caaaagttcg   20340

atttattcaa caaagccacg ttgtgtctca aaatctctga tgttacattg cacaagataa   20400

aaatatatca tcatgaacaa taaaactgtc tgcttacata aacagtaata caaggggtgt   20460

tatgagccat attcaacggg aaacgtcttg ctcgactcta gagctcgttc ctcgaggcct   20520

cgaggcctcg aggaacggta cctgcgggga agcttacaat aatgtgtgtt gttaagtctt   20580

gttgcctgtc atcgtctgac tgactttcgt cataaatccc ggcctccgta acccagcttt   20640

gggcaagctc acggatttga tccggcggaa cgggaatatc gagatgccgg gctgaacgct   20700

gcagttccag ctttcccttt cgggacaggt actccagctg attgattatc tgctgaaggg   20760

tcttggttcc acctcctggc acaatgcgaa tgattacttg agcgcgatcg ggcatccaat   20820
```

-continued

```
tttctcccgt caggtgcgtg gtcaagtgct acaaggcacc tttcagtaac gagcgaccgt    20880

cgatccgtcg ccgggatacg gacaaaatgg agcgcagtag tccatcgagg gcggcgaaag    20940

cctcgccaaa agcaatacgt tcatctcgca cagcctccag atccgatcga gggtcttcgg    21000

cgtaggcaga tagaagcatg gatacattgc ttgagagtat tccgatggac tgaagtatgg    21060

cttccatctt ttctcgtgtg tctgcatcta tttcgagaaa gcccccgatg cggcgcaccg    21120

caacgcgaat tgccatacta tccgaaagtc ccagcaggcg cgcttgatag gaaaaggttt    21180

catactcggc cgatcgcaga cgggcactca cgaccttgaa cccttcaact ttcagggatc    21240

gatgctggtt gatggtagtc tcactcgacg tggctctggt gtgttttgac atagcttcct    21300

ccaaagaaag cggaaggtct ggatactcca gcacgaaatg tgcccgggta gacggatgga    21360

agtctagccc tgctcaatat gaaatcaaca gtacatttac agtcaatact gaatatactt    21420

gctacatttg caattgtctt ataacgaatg tgaaataaaa atagtgtaac aacgctttta    21480

ctcatcgata atcacaaaaa catttatacg aacaaaaata caaatgcact ccggtttcac    21540

aggataggcg ggatcagaat atgcaacttt tgacgttttg ttctttcaaa gggggtgctg    21600

gcaaaaccac cgcactcatg ggcctttgcg ctgctttggc aaatgacggt aaacgagtgg    21660

ccctctttga tgccgacgaa aaccggcctc tgacgcgatg gagagaaaac gccttacaaa    21720

gcagtactgg gatcctcgct gtgaagtcta ttccgccgac gaaatgcccc ttcttgaagc    21780

agcctatgaa aatgccgagc tcgaaggatt tgattatgcg ttggccgata cgcgtggcgg    21840

ctcgagcgag ctcaacaaca caatcatcgc tagctcaaac ctgcttctga tccccaccat    21900

gctaacgccg ctcgacatcg atgaggcact atctacctac cgctacgtca tcgagctgct    21960

gttgagtgaa aatttggcaa ttcctacagc tgtttttgcgc caacgcgtcc cggtcggccg    22020

attgacaaca tcgcaacgca ggatgtcaga gacgctagag agccttccag ttgtaccgtc    22080

tcccatgcat gaaagagatg catttgccgc gatgaaagaa cgcggcatgt tgcatcttac    22140

attactaaac acgggaactg atccgacgat gcgcctcata gagaggaatc ttcggattgc    22200

gatggaggaa gtcgtggtca tttcgaaact gatcagcaaa atcttggagg cttgaagatg    22260

gcaattcgca agcccgcatt gtcggtcggc gaagcacggc ggcttgctgg tgctcgaccc    22320

gagatccacc atcccaaccc gacacttgtt ccccagaagc tggacctcca gcacttgcct    22380

gaaaaagccg acgagaaaga ccagcaacgt gagcctctcg tcgccgatca catttacagt    22440

cccgatcgac aacttaagct aactgtggat gcccttagtc cacctccgtc cccgaaaaag    22500

ctccaggttt ttctttcagc gcgaccgccc gcgcctcaag tgtcgaaaac atatgacaac    22560

ctcgttcggc aatacagtcc ctcgaagtcg ctacaaatga ttttaaggcg cgcgttggac    22620

gatttcgaaa gcatgctggc agatggatca tttcgcgtgg ccccgaaaag ttatccgatc    22680

ccttcaacta cagaaaaatc cgttctcgtt cagacctcac gcatgttccc ggttgcgttg    22740

ctcgaggtcg ctcgaagtca tttttgatccg ttggggttgg agaccgctcg agctttcggc    22800

cacaagctgg ctaccgccgc gctcgcgtca ttctttgctg gagagaagcc atcgagcaat    22860

tggtgaagag ggacctatcg gaacccctca ccaaatattg agtgtaggtt tgaggccgct    22920

ggccgcgtcc tcagtcacct tttgagccag ataattaaga gccaaatgca attggctcag    22980

gctgccatcg tcccccgtg cgaaacctgc acgtccgcgt caaagaaata accggcacct    23040

cttgctgttt ttatcagttg agggcttgac ggatccgcct caagtttgcg gcgcagccgc    23100
```

-continued

```
aaaatgagaa catctatact cctgtcgtaa acctcctcgt cgcgtactcg actggcaatg   23160

agaagttgct cgcgcgatag aacgtcgcgg ggtttctcta aaaacgcgag gagaagattg   23220

aactcacctg ccgtaagttt cacctcaccg ccagcttcgg acatcaagcg acgttgcctg   23280

agattaagtg tccagtcagt aaaacaaaaa gaccgtcggt ctttggagcg gacaacgttg   23340

gggcgcacgc gcaaggcaac ccgaatgcgt gcaagaaact ctctcgtact aaacggctta   23400

gcgataaaat cacttgctcc tagctcgagt gcaacaactt tatccgtctc ctcaaggcgg   23460

tcgccactga taattatgat tggaatatca gactttgccg ccagatttcg aacgatctca   23520

agcccatctt cacgacctaa atttagatca acaaccacga catcgaccgt cgcggaagag   23580

agtactctag tgaactgggt gctgtcggct accgcggtca ctttgaaggc gtggatcgta   23640

aggtattcga taataagatg ccgcatagcg acatcgtcat cgataagaag aacgtgtttc   23700

aacggctcac ctttcaatct aaaatctgaa cccttgttca cagcgcttga gaaattttca   23760

cgtgaaggat gtacaatcat ctccagctaa atgggcagtt cgtcagaatt gcggctgacc   23820

gcggatgacg aaaatgcgaa ccaagtattt caattttatg acaaaagttc tcaatcgttg   23880

ttacaagtga aacgcttcga ggttacagct actattgatt aaggagatcg cctatggtct   23940

cgccccggcg tcgtgcgtcc gccgcgagcc agatctcgcc tacttcataa acgtcctcat   24000

aggcacggaa tggaatgatg acatcgatcg ccgtagagag catgtcaatc agtgtgcgat   24060

cttccaagct agcaccttgg gcgctacttt tgacaaggga aaacagtttc ttgaatcctt   24120

ggattggatt cgcgccgtgt attgttgaaa tcgatcccgg atgtcccgag acgacttcac   24180

tcagataagc ccatgctgca tcgtcgcgca tctcgccaag caatatccgg tccggccgca   24240

tacgcagact tgcttggagc aagtgctcgg cgctcacagc acccagccca gcaccgttct   24300

tggagtagag tagtctaaca tgattatcgt gtggaatgac gagttcgagc gtatcttcta   24360

tggtgattag cctttcctgg gggggatgg cgctgatcaa ggtcttgctc attgttgtct   24420

tgccgcttcc ggtagggcca catagcaaca tcgtcagtcg gctgacgacg catgcgtgca   24480

gaaacgcttc caaatccccg ttgtcaaaat gctgaaggat agcttcatca tcctgatttt   24540

ggcgtttcct tcgtgtctgc cactggttcc acctcgaagc atcataacgg gaggagactt   24600

ctttaagacc agaaacacgc gagcttggcc gtcgaatggt caagctgacg gtgcccgagg   24660

gaacggtcgg cggcagacag atttgtagtc gttcaccacc aggaagttca gtggcgcaga   24720

gggggttacg tggtccgaca tcctgctttc tcagcgcgcc cgctaaaata gcgatatctt   24780

caagatcatc ataagagacg ggcaaaggca tcttggtaaa aatgccggct tggcgcacaa   24840

atgcctctcc aggtcgattg atcgcaattt cttcagtctt cgggtcatcg agccattcca   24900

aaatcggctt cagaagaaag cgtagttgcg gatccacttc catttacaat gtatcctatc   24960

tctaagcgga aatttgaatt cattaagagc ggcggttcct ccccgcgtg gcgccgccag   25020

tcaggcggag ctggtaaaca ccaaagaaat cgaggtcccg tgctacgaaa atggaaacgg   25080

tgtcaccctg attcttcttc agggttggcg gtatgttgat ggttgcctta agggctgtct   25140

cagttgtctg ctcaccgtta ttttgaaagc tgttgaagct catcccgcca cccgagctgc   25200

cggcgtaggt gctagctgcc tggaaggcgc cttgaacaac actcaagagc atagctccgc   25260

taaaacgctg ccagaagtgg ctgtcgaccg agcccggcaa tcctgagcga ccgagttcgt   25320

ccgcgcttgg cgatgttaac gagatcatcg catggtcagg tgtctcggcg cgatcccaca   25380
```

-continued

```
acacaaaaac gcgcccatct ccctgttgca agccacgctg tatttcgcca acaacggtgg   25440

tgccacgatc aagaagcacg atattgttcg ttgttccacg aatatcctga ggcaagacac   25500

actttacata gcctgccaaa tttgtgtcga ttgcggtttg caagatgcac ggaattattg   25560

tcccttgcgt taccataaaa tcggggtgcg gcaagagcgt ggcgctgctg ggctgcagct   25620

cggtgggttt catacgtatc gacaaatcgt tctcgccgga cacttcgcca ttcggcaagg   25680

agttgtcgtc acgcttgcct tcttgtcttc ggcccgtgtc gccctgaatg gcgcgtttgc   25740

tgaccccttg atcgccgctg ctatatgcaa aaatcggtgt ttcttccggc cgtggctcat   25800

gccgctccgg ttcgcccctc ggcggtagag gagcagcagg ctgaacagcc tcttgaaccg   25860

ctggaggatc cggcggcacc tcaatcggag ctggatgaaa tggcttggtg tttgttgcga   25920

tcaaagttga cggcgatgcg ttctcattca ccttcttttg gcgcccacct agccaaatga   25980

ggcttaatga taacgcgaga acgacacctc cgacgatcaa tttctgagac cccgaaagac   26040

gccggcgatg tttgtcggag accagggatc cagatgcatc aacctcatgt gccgcttgct   26100

gactatcgtt attcatccct tcgccccctt caggacgcgt ttcacatcgg gcctcaccgt   26160

gcccgtttgc ggcctttggc caacgggatc gtaagcggtg ttccagatac atagtactgt   26220

gtggccatcc ctcagacgcc aacctcggga aaccgaagaa atctcgacat cgctcccttt   26280

aactgaatag ttggcaacag cttccttgcc atcaggattg atggtgtaga tggagggtat   26340

gcgtacattg cccggaaagt ggaataccgt cgtaaatcca ttgtcgaaga cttcgagtgg   26400

caacagcgaa cgatcgcctt gggcgacgta gtgccaatta ctgtccgccg caccaagggc   26460

tgtgacaggc tgatccaata aattctcagc tttccgttga tattgtgctt ccgcgtgtag   26520

tctgtccaca acagccttct gttgtgcctc ccttcgccga gccgccgcat cgtcggcggg   26580

gtaggcgaat tggacgctgt aatagagatc gggctgctct ttatcgaggt gggacagagt   26640

cttggaactt atactgaaaa cataacggcg catcccggag tcgcttgcgg ttagcacgat   26700

tactggctga ggcgtgagga cctggcttgc cttgaaaaat agataatttc cccgcggtag   26760

ggctgctaga tctttgctat ttgaaacggc aaccgctgtc accgtttcgt tcgtggcgaa   26820

tgttacgacc aaagtagctc caaccgccgt cgagaggcgc accacttgat cgggattgta   26880

agccaaataa cgcatgcgcg gatctagctt gcccgccatt ggagtgtctt cagcctccgc   26940

accagtcgca gcggcaaata aacatgctaa aatgaaaagt gcttttctga tcatggttcg   27000

ctgtggccta cgtttgaaac ggtatcttcc gatgtctgat aggaggtgac aaccagacct   27060

gccgggttgg ttagtctcaa tctgccgggc aagctggtca ccttttcgta gcgaactgtc   27120

gcggtccacg tactcaccac aggcattttg ccgtcaacga cgagggtcct tttatagcga   27180

atttgctgcg tgcttggagt tacatcattt gaagcgatgt gctcgacctc caccctgccg   27240

cgtttgccaa gaatgacttg aggcgaactg ggattgggat agttgaagaa ttgctggtaa   27300

tcctggcgca ctgttggggc actgaagttc gataccaggt cgtaggcgta ctgagcggtg   27360

tcggcatcat aactctcgcg caggcgaacg tactcccaca atgaggcgtt aacgacggcc   27420

tcctcttgag ttgcaggcaa tcgcgagaca gacacctcgc tgtcaacggt gccgtccggc   27480

cgtatccata gatatacggg cacaagcctg ctcaacggca ccattgtggc tatagcgaac   27540

gcttgagcaa catttcccaa aatcgcgata gctgcgacag ctgcaatgag tttggagaga   27600

cgtcgcgccg atttcgctcg cgcggtttga aaggcttcta cttccttata gtgctcggca   27660
```

-continued

```
aggctttcgc gcgccactag catggcatat tcaggccccg tcatagcgtc cacccgaatt   27720

gccgagctga agatctgacg gagtaggctg ccatcgcccc acattcagcg ggaagatcgg   27780

gcctttgcag ctcgctaatg tgtcgtttgt ctggcagccg ctcaaagcga caactaggca   27840

cagcaggcaa tacttcatag aattctccat tgaggcgaat ttttgcgcga cctagcctcg   27900

ctcaacctga gcgaagcgac ggtacaagct gctggcagat tgggttgcgc cgctccagta   27960

actgcctcca atgttgccgg cgatcgccgg caaagcgaca atgagcgcat cccctgtcag   28020

aaaaaacata tcgagttcgt aaagaccaat gatcttggcc gcggtcgtac cggcgaaggt   28080

gattacacca agcataaggg tgagcgcagt cgcttcggtt aggatgacga tcgttgccac   28140

gaggtttaag aggagaagca agagaccgta ggtgataagt tgcccgatcc acttagctgc   28200

gatgtcccgc gtgcgatcaa aaatatatcc gacgaggatc agaggcccga tcgcgagaag   28260

cactttcgtg agaattccaa cggcgtcgta aactccgaag gcagaccaga gcgtgccgta   28320

aaggacccac tgtgcccctt ggaaagcaag gatgtcctgg tcgttcatcg gaccgatttc   28380

ggatgcgatt ttctgaaaaa cggcctgggt cacggcgaac attgtatcca actgtgccgg   28440

aacagtctgc agaggcaagc cggttacact aaactgctga acaaagtttg ggaccgtctt   28500

ttcgaagatg gaaaccacat agtcttggta gttagcctgc ccaacaatta gagcaacaac   28560

gatggtgacc gtgatcaccc gagtgatacc gctacgggta tcgacttcgc cgcgtatgac   28620

taaaataccc tgaacaataa tccaaagagt gacacaggcg atcaatggcg cactcaccgc   28680

ctcctggata gtctcaagca tcgagtccaa gcctgtcgtg aaggctacat cgaagatcgt   28740

atgaatggcc gtaaacggcg ccggaatcgt gaaattcatc gattggacct gaacttgact   28800

ggtttgtcgc ataatgttgg ataaaatgag ctcgcattcg gcgaggatgc gggcggatga   28860

acaaatcgcc cagccttagg ggagggcacc aaagatgaca gcggtctttt gatgctcctt   28920

gcgttgagcg gccgcctctt ccgcctcgtg aaggccggcc tgcgcggtag tcatcgttaa   28980

taggcttgtc gcctgtacat tttgaatcat tgcgtcatgg atctgcttga gaagcaaacc   29040

attggtcacg gttgcctgca tgatattgcg agatcgggaa agctgagcag acgtatcagc   29100

attcgccgtc aagcgtttgt ccatcgtttc cagattgtca gccgcaatgc cagcgctgtt   29160

tgcggaaccg gtgatctgcg atcgcaacag gtccgcttca gcatcactac ccacgactgc   29220

acgatctgta tcgctggtga tcgcacgtgc cgtggtcgac attggcattc gcggcgaaaa   29280

catttcattg tctaggtcct tcgtcgaagg atactgattt ttctggttga gcgaagtcag   29340

tagtccagta acgccgtagg ccgacgtcaa catcgtaacc atcgctatag tctgagtgag   29400

attctccgca gtcgcgagcg cagtcgcgag cgtctcagcc tccgttgccg ggtcgctaac   29460

aacaaactgc gcccgcgcgg ctgaatata tagaaagctg caggtcaaaa ctgttgcaat   29520

aagttgcgtc gtcttcatcg tttcctacct tatcaatctt ctgcctcgtg gtgacgggcc   29580

atgaattcgc tgagccagcc agatgagttg ccttcttgtg cctcgcgtag tcgagttgca   29640

aagcgcaccg tgttggcacg ccccgaaagc acggcgacat attcacgcat atcccgcaga   29700

tcaaattcgc agatgacgct tccactttct cgtttaagaa gaaacttacg gctgccgacc   29760

gtcatgtctt cacggatcgc ctgaaattcc ttttcggtac atttcagtcc atcgacataa   29820

gccgatcgat ctgcggttgg tgatggatag aaaatcttcg tcatacattg cgcaaccaag   29880

ctggctccta gcggcgattc cagaacatgc tctggttgct gcgttgccag tattagcatc   29940
```

-continued

```
ccgttgtttt ttcgaacggt caggaggaat ttgtcgacga cagtcgaaaa tttagggttt   30000

aacaaatagg cgcgaaactc atcgcagctc atcacaaaac ggcggccgtc gatcatggct   30060

ccaatccgat gcaggagata tgctgcagcg ggagcgcata cttcctcgta ttcgagaaga   30120

tgcgtcatgt cgaagccggt aatcgacgga tctaacttta cttcgtcaac ttcgccgtca   30180

aatgcccagc caagcgcatg gccccggcac cagcgttgga gccgcgctcc tgcgccttcg   30240

gcgggcccat gcaacaaaaa ttcacgtaac cccgcgattg aacgcatttg tggatcaaac   30300

gagagctgac gatggatacc acggaccaga cggcggttct cttccggaga aatcccaccc   30360

cgaccatcac tctcgatgag agccacgatc cattcgcgca gaaaatcgtg tgaggctgct   30420

gtgttttcta ggccacgcaa cggcgccaac ccgctgggtg tgcctctgtg aagtgccaaa   30480

tatgttcctc ctgtggcgcg aaccagcaat tcgccacccc ggtccttgtc aaagaacacg   30540

accgtacctg cacggtcgac catgctctgt tcgagcatgg ctagaacaaa catcatgagc   30600

gtcgtcttac ccctcccgat aggcccgaat attgccgtca tgccaacatc gtgctcatgc   30660

gggatatagt cgaaaggcgt tccgccattg gtacgaaatc gggcaatcgc gttgccccag   30720

tggcctgagc tggcgccctc tggaaagttt tcgaaagaga caaaccctgc gaaattgcgt   30780

gaagtgattg cgccagggcg tgtgcgccac ttaaaattcc ccggcaattg ggaccaatag   30840

gccgcttcca taccaatacc ttcttggaca accacggcac ctgcatccgc cattcgtgtc   30900

cgagcccgcg cgccctgtc cccaagacta ttgagatcgt ctgcatagac gcaaaggctc   30960

aaatgatgtg agcccataac gaattcgttg ctcgcaagtg cgtcctcagc ctcggataat   31020

ttgccgattt gagtcacggc tttatcgccg gaactcagca tctggctcga tttgaggcta   31080

agtttcgcgt gcgcttgcgg gcgagtcagg aacgaaaaac tctgcgtgag aacaagtgga   31140

aaatcgaggg atagcagcgc gttgagcatg cccggccgtg tttttgcagg gtattcgcga   31200

aacgaataga tggatccaac gtaactgtct tttggcgttc tgatctcgag tcctcgcttg   31260

ccgcaaatga ctctgtcggt ataaatcgaa gcgccgagtg agccgctgac gaccggaacc   31320

ggtgtgaacc gaccagtcat gatcaaccgt agcgcttcgc caatttcggt gaagagcaca   31380

ccctgcttct cgcggatgcc aagacgatgc aggccatacg ctttaagaga gccagcgaca   31440

acatgccaaa gatcttccat gttcctgatc tggcccgtga gatcgttttc ccttttttccg   31500

cttagcttgg tgaacctcct ctttaccttc cctaaagccg cctgtgggta gacaatcaac   31560

gtaaggaagt gttcattgcg gaggagttgg ccggagagca cgcgctgttc aaaagcttcg   31620

ttcaggctag cggcgaaaac actacggaag tgtcgcggcg ccgatgatgg cacgtcggca   31680

tgacgtacga ggtgagcata tattgacaca tgatcatcag cgatattgcg caacagcgtg   31740

ttgaacgcac gacaacgcgc attgcgcatt tcagtttcct caagctcgaa tgcaacgcca   31800

tcaattctcg caatggtcat gatcgatccg tcttcaagaa ggacgatatg gtcgctgagg   31860

tggccaatat aagggagata gatctcaccg gatctttcgg tcgttccact cgcgccgagc   31920

atcacaccat tcctctccct cgtgggggaa ccctaattgg atttgggcta acagtagcgc   31980

cccccaaac tgcactatca atgcttcttc ccgcggtccg caaaaatagc aggacgacgc   32040

tcgccgcatt gtagtctcgc tccacgatga gccgggctgc aaaccataac ggcacgagaa   32100

cgacttcgta gagcgggttc tgaacgataa cgatgacaaa gccggcgaac atcatgaata   32160

accctgccaa tgtcagtggc accccaagaa acaatgcggg ccgtgtggct gcgaggtaaa   32220
```

-continued

```
gggtcgattc ttccaaacga tcagccatca actaccgcca gtgagcgttt ggccgaggaa   32280

gctcgcccca aacatgataa caatgccgcc gacgacgccg gcaaccagcc caagcgaagc   32340

ccgcccgaac atccaggaga tcccgatagc gacaatgccg agaacagcga gtgactggcc   32400

gaacggacca aggataaacg tgcatatatt gttaaccatt gtggcggggt cagtgccgcc   32460

acccgcagat tgcgctgcgg cgggtccgga tgaggaaatg ctccatgcaa ttgcaccgca   32520

caagcttggg gcgcagctcg atatcacgcg catcatcgca ttcgagagcg agaggcgatt   32580

tagatgtaaa cggtatctct caaagcatcg catcaatgcg cacctcctta gtataagtcg   32640

aataagactt gattgtcgtc tgcggatttg ccgttgtcct ggtgtggcgg tggcggagcg   32700

attaaaccgc cagcgccatc ctcctgcgag cggcgctgat atgacccca aacatcccac    32760

gtctcttcgg attttagcgc ctcgtgatcg tcttttggag gctcgattaa cgcgggcacc   32820

agcgattgag cagctgtttc aacttttcgc acgtagccgt ttgcaaaacc gccgatgaaa   32880

ttaccggtgt tgtaagcgga gatcgcccga cgaagcgcaa attgcttctc gtcaatcgtt   32940

tcgccgcctg cataacgact tttcagcatg tttgcagcgg cagataatga tgtgcacgcc   33000

tggagcgcac cgtcaggtgt cagaccgagc atagaaaaat ttcgagagtt tatttgcatg   33060

aggccaacat ccagcgaatg ccgtgcatcg agacggtgcc tgacgacttg ggttgcttgg   33120

ctgtgatctt gccagtgaag cgtttcgccg gtcgtgttgt catgaatcgc taaaggatca   33180

aagcgactct ccaccttagc tatcgccgca agcgtagatg tcgcaactga tggggcacac   33240

ttgcgagcaa catggtcaaa ctcagcagat gagagtggcg tggcaaggct cgacgaacag   33300

aaggagacca tcaaggcaag agaaagcgac cccgatctct taagcatacc ttatctcctt   33360

agctcgcaac taacaccgcc tctcccgttg gaagaagtgc gttgttttat gttgaagatt   33420

atcgggaggg tcggttactc gaaaatttc aattgcttct ttatgatttc aattgaagcg   33480

agaaacctcg cccggcgtct tggaacgcaa catggaccga gaaccgcgca tccatgacta   33540

agcaaccgga tcgacctatt caggccgcag ttggtcaggt caggctcaga acgaaaatgc   33600

tcggcgaggt tacgctgtct gtaaacccat tcgatgaacg ggaagcttcc ttccgattgc   33660

tcttggcagg aatattggcc catgcctgct tgcgctttgc aaatgctctt atcgcgttgg   33720

tatcatatgc cttgtccgcc agcagaaacg cactctaagc gattatttgt aaaaatgttt   33780

cggtcatgcg gcggtcatgg gcttgacccg ctgtcagcgc aagacggatc ggtcaaccgt   33840

cggcatcgac aacagcgtga atcttggtgg tcaaaccgcc acgggaacgt cccatacagc   33900

catcgtcttg atcccgctgt ttcccgtcgc cgcatgttgg tggacgcgga cacaggaact   33960

gtcaatcatg acgacattct atcgaaagcc ttggaaatca cactcagaat atgatcccag   34020

acgtctgcct cacgccatcg tacaaagcga ttgtagcagg ttgtacagga accgtatcga   34080

tcaggaacgt ctgcccaggg cgggcccgtc cggaagcgcc acaagatgac attgatcacc   34140

cgcgtcaacg cgcggcacgc gacgcggctt atttgggaac aaaggactga acaacagtcc   34200

attcgaaatc ggtgacatca aagcggggac gggttatcag tggcctccaa gtcaagcctc   34260

aatgaatcaa aatcagaccg atttgcaaac ctgatttatg agtgtgcggc ctaaatgatg   34320

aaatcgtcct tctagatcgc ctccgtggtg tagcaacacc tcgcagtatc gccgtgctga   34380

ccttggccag ggaattgact ggcaagggtg ctttcacatg accgctcttt ggccgcgat    34440

agatgatttc gttgctgctt tgggcacgta gaaggagaga agtcatatcg gagaaattcc   34500
```

-continued

```
tcctggcgcg agagcctgct ctatcgcgac ggcatcccac tgtcgggaac agaccggatc   34560

attcacgagg cgaaagtcgt caacacatgc gttataggca tcttcccttg aaggatgatc   34620

ttgttgctgc caatctggag gtgcggcagc cgcaggcaga tgcgatctca gcgcaacttg   34680

cggcaaaaca tctcactcac ctgaaaacca ctagcgagtc tcgcgatcag acgaaggcct   34740

tttacttaac gacacaatat ccgatgtctg catcacaggc gtcgctatcc cagtcaatac   34800

taaagcggtg caggaactaa agattactga tgacttaggc gtgccacgag gcctgagacg   34860

acgcgcgtag acagtttttt gaaatcatta tcaaagtgat ggcctccgct gaagcctatc   34920

acctctcgcgc cggtctgtcg gagagatggg caagcattat tacggtcttc gcgcccgtac   34980

atgcattgga cgattgcagg gtcaatggat ctgagatcat ccagaggatt gccgccctta   35040

ccttccgttt cgagttggag ccagccccta aatgagacga catagtcgac ttgatgtgac   35100

aatgccaaga gagagatttg cttaacccga ttttttttgct caagcgtaag cctattgaag   35160

cttgccggca tgacgtccgc gccgaaagaa tatcctacaa gtaaaacatt ctgcacaccg   35220

aaatgcttgg tgtagacatc gattatgtga ccaagatcct tagcagtttc gcttggggac   35280

cgctccgacc agaaataccg aagtgaactg acgccaatga caggaatccc ttccgtctgc   35340

agataggtac catcgataga tctgctgcct cgcgcgtttc ggtgatgacg gtgaaaacct   35400

ctgacacatg cagctcccgg agacggtcac agcttgtctg taagcggatg ccgggagcag   35460

acaagcccgt cagggcgcgt cagcgggtgt tggcgggtgt cggggcgcag ccatgaccca   35520

gtcacgtagc gatagcggag tgtatactgg cttaactatg cggcatcaga gcagattgta   35580

ctgagagtgc accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc   35640

atcaggcgct cttccgcttc ctcgctcact gactcgctgc gctcggtcgt tcggctgcgg   35700

cgagcggtat cagctcactc aaaggcggta atacggttat ccacagaatc agggggataac   35760

gcaggaaaga acatgtgagc aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg   35820

ttgctggcgt ttttccatag gctccgcccc cctgacgagc atcacaaaaa tcgacgctca   35880

agtcagaggt ggcgaaaccc gacaggacta taaagatacc aggcgtttcc ccctggaagc   35940

tccctcgtgc gctctcctgt ccgaccctg ccgcttaccg gatacctgtc cgcctttctc   36000

ccttcgggaa gcgtggcgct ttctcatagc tcacgctgta ggtatctcag ttcggtgtag   36060

gtcgttcgct ccaagctggg ctgtgtgcac gaaccccccg ttcagcccga ccgctgcgcc   36120

ttatccggta actatcgtct tgagtccaac ccggtaagac acgacttatc gccactggca   36180

gcagccactg gtaacaggat tagcagagcg aggtatgtag gcggtgctac agagttcttg   36240

aagtggtggc ctaactacgg ctacactaga aggacagtat ttggtatctg cgctctgctg   36300

aagccagtta ccttcggaaa aagagttggt agctcttgat ccggcaaaca aaccaccgct   36360

ggtagcggtg gtttttttgt ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa   36420

gaagatcctt tgatctttc tacggggtct gacgctcagt ggaacgaaaa ctcacgttaa   36480

gggattttgg tcatgagatt atcaaaaagg atcttcacct agatcctttt aaattaaaaa   36540

tgaagtttta aatcaatcta agtatatat gagtaaactt ggtctgacag ttaccaatgc   36600

ttaatcagtg aggcacctat ctcagcgatc tgtctatttc gttcatccat agttgcctga   36660

ctccccgtcg tgtagataac tacgatacgg gagggcttac catctggccc cagtgctgca   36720

atgataccgc gagacccacg ctcaccggct ccagatttat cagcaataaa ccagccagcc   36780
```

-continued

```
ggaagggccg agcgcagaag tggtcctgca actttatccg cctccatcca gtctattaat   36840

tgttgccggg aagctagagt aagtagttcg ccagttaata gtttgcgcaa cgttgttgcc   36900

attgctgcag gggggggggg gggggggggac ttccattgtt cattccacgg acaaaaacag   36960

agaaaggaaa cgacagaggc caaaaagcct cgctttcagc acctgtcgtt tcctttcttt   37020

tcagagggta ttttaaataa aaacattaag ttatgacgaa gaagaacgga aacgccttaa   37080

accggaaaat tttcataaat agcgaaaacc cgcgaggtcg ccgccccgta acctgtcgga   37140

tcaccggaaa ggacccgtaa agtgataatg attatcatct acatatcaca acgtgcgtgg   37200

aggccatcaa accacgtcaa ataatcaatt atgacgcagg tatcgtatta attgatctgc   37260

atcaacttaa cgtaaaaaca acttcagaca atacaaatca gcgacactga atacggggca   37320

acctcatgtc cccccccccc cccccctgc aggcatcgtg gtgtcacgct cgtcgtttgg   37380

tatggcttca ttcagctccg gttcccaacg atcaaggcga gttacatgat cccccatgtt   37440

gtgcaaaaaa gcggttagct ccttcggtcc tccgatcgtt gtcagaagta agttggccgc   37500

agtgttatca ctcatggtta tggcagcact gcataattct cttactgtca tgccatccgt   37560

aagatgcttt tctgtgactg gtgagtactc aaccaagtca ttctgagaat agtgtatgcg   37620

gcgaccgagt tgctcttgcc cggcgtcaac acgggataat accgcgccac atagcagaac   37680

tttaaaagtg ctcatcattg gaaaacgttc ttcggggcga aaactctcaa ggatcttacc   37740

gctgttgaga tccagttcga tgtaacccac tcgtgcaccc aactgatctt cagcatcttt   37800

tactttcacc agcgtttctg ggtgagcaaa aacaggaagg caaaatgccg caaaaaaggg   37860

aataagggcg acacggaaat gttgaatact catactcttc cttttcaat attattgaag   37920

catttatcag ggttattgtc tcatgagcgg atacatattt gaatgtattt agaaaaataa   37980

acaaataggg gttccgcgca catttccccg aaaagtgcca cctgacgtct aagaaaccat   38040

tattatcatg acattaacct ataaaaatag gcgtatcacg aggccctttc gtcttcaaga   38100

attggtcgac gatcttgctg cgttcggata ttttcgtgga gttcccgcca cagacccgga   38160

ttgaaggcga gatccagcaa ctcgcgccag atcatcctgt gacggaactt ggcgcgtga   38220

tgactggcca ggacgtcggc cgaaagagcg acaagcagat cacgctttc gacagcgtcg   38280

gatttgcgat cgaggatttt tcggcgctgc gctacgtccg cgaccgcgtt gagggatcaa   38340

gccacagcag cccactcgac cttctagccg acccagacga gccaagggat cttttttggaa   38400

tgctgctccg tcgtcaggct ttccgacgtt tgggtggttg aacagaagtc attatcgtac   38460

ggaatgccaa gcactcccga ggggaaccct gtggttggca tgcacataca aatggacgaa   38520

cggataaacc tttttcacgcc cttttaaata tccgttattc taataaacgc tcttttctct   38580

taggtttacc cgccaatata tcctgtcaaa cactgatagt ttaaactgaa ggcgggaaac   38640

gacaatctga tcatgagcgg agaattaagg gagtcacgtt atgacccccg ccgatgacgc   38700

gggacaagcc gttttacgtt tggaactgac agaaccgcaa cgttgaagga gccactcagc   38760

aagctggtac gattgtaata cgactcacta tagggcgaat tgagcgctgt ttaaacgctc   38820

ttcaactgga agagcggtta cccggaccga agcttgaagt tcctattccg aagttcctat   38880

tctctagaaa gtataggaac ttcagatctc gatgctcacc ctgttgtttg gtgttacttc   38940

tgcaggtcga ctctagagga tccaccatga gcccagaacg acgcccggcc gacatccgcc   39000

gtgccaccga ggcggacatg ccggcggtct gcaccatcgt caaccactac atcgagacaa   39060
```

-continued

```
gcacggtcaa cttccgtacc gagccgcagg aaccgcagga ctggacggac gacctcgtcc   39120

gtctgcggga gcgctatccc tggctcgtcg ccgaggtgga cggcgaggtc gccggcatcg   39180

cctacgcggg ccctggaag gcacgcaacg cctacgactg gacggccgag tcgaccgtgt   39240

acgtctcccc ccgccaccag cggacgggac tgggctccac gctctacacc cacctgctga   39300

agtccctgga ggcacagggc ttcaagagcg tggtcgctgt catcgggctg cccaacgacc   39360

cgagcgtgcg catgcacgag gcgctcggat atgcccccg cggcatgctg cgggcggccg   39420

gcttcaagca cgggaactgg catgacgtgg gtttctggca gctggacttc agcctgccgg   39480

taccgccccg tccggtcctg cccgtcaccg agatctgatc cgtcgaccaa cctagacttg   39540

tccatcttct ggattggcca acttaattaa tgtatgaaat aaaaggatgc acacatagtg   39600

acatgctaat cactataatg tgggcatcaa agttgtgtgt tatgtgtaat tactagttat   39660

ctgaataaaa gagaaagaga tcatccatat ttcttatcct aaatgaatgt cacgtgtctt   39720

tataattctt tgatgaacca gatgcatttc attaaccaaa tccatataca tataaatatt   39780

aatcatatat aattaatatc aattgggtta gcaaacaaa tctagtctag gtgtgttttg   39840

cgaattgcgg ccgcgatctg gggaattccc atggacaccg gtaattccca tgatcttctc   39900

tccttcatca atggatgcca tgtttcataa caataacacc aaatgtttga tgagctacca   39960

acaattgcgc aaagactatg gctaagctcg agctcgctcg ctacaagttg ttgactttca   40020

aatacaagtt tgttttggga acaccaaata ttctacatga tctttcacta agttgcgcac   40080

cactatcaaa agattatcta ggccattatt caagtaaaga gtgaacacgt ctaagaccca   40140

caaccacacc aaatagaata cgcatacatg caacatattg tgcaagaagt atccaactgg   40200

actccatgt attctaaaac tattttcgta gagttaaagt tatgacaaac ttatcaaata   40260

aaaatttgaa cgctggacca aaactttcat cttttcaaatc caccatcgtc tatcctcata   40320

aattgttttg attataacac atctacgtaa atcatttgtt ttgaacaata ctaatttaat   40380

tttattaagt caaataacct gcttagaaaa taatccctcc acctcattta acaatttctt   40440

gtcaaacaca caccaagaaa aaaattaatg aaagagaaaa gaaatgaaaa ggacatggag   40500

ttgaatacta gcaaaattga ttgaaggaag attcacaatt gaaattgaaa ccatttaatt   40560

tattttcggg tccataataa taaattggta agaataaaaa cccgatcaag tccggtacag   40620

tacaattcca ctccaccaac tccttactta aaccccctatt tatacccact ctcatcctca   40680

ctcttccttc acctctcaca ctctcttctc tctctcaaaa ccctcacaca aacgctgcgt   40740

ttagtgtaag aaattcaatc cggcgccttg gcgcgccgat catccacaag tttgtacaaa   40800

aaagctgaac gagaaacgta aaatgatata aatatcaata tattaaatta gattttgcat   40860

aaaaaacaga ctacataata ctgtaaaaca caacatatcc agtcactatg gcggccgcat   40920

taggcacccc aggctttaca ctttatgctt ccggctcgta taatgtgtgg attttgagtt   40980

aggatttaaa tacgcgttga tccggcttac taaaagccag ataacagtat gcgtatttgc   41040

gcgctgattt ttgcggtata agaatatata ctgatatgta tacccgaagt atgtcaaaaa   41100

gaggtatgct atgaagcagc gtattacagt gacagttgac agcgacagct atcagttgct   41160

caaggcatat atgatgtcaa tatctccggt ctggtaagca caaccatgca gaatgaagcc   41220

cgtcgtctgc gtgccgaacg ctggaaagcg gaaaatcagg aagggatggc tgaggtcgcc   41280

cggtttattg aaatgaacgg ctcttttgct gacgagaaca ggggctggtg aaatgcagtt   41340
```

-continued

```
taaggtttac acctataaaa gagagagccg ttatcgtctg tttgtggatg tacagagtga   41400

tatcattgac acgcccggtc gacggatggt gatcccctg gccagtgcac gtctgctgtc    41460

agataaagtc tcccgtgaac tttacccggt ggtgcatatc ggggatgaaa gctggcgcat   41520

gatgaccacc gatatggcca gtgtgccggt ctccgttatc ggggaagaag tggctgatct   41580

cagccaccgc gaaaatgaca tcaaaaacgc cattaacctg atgttctggg gaatataaat   41640

gtcaggctcc cttatacaca gccagtctgc aggtcgacca tagtgactgg atatgttgtg   41700

ttttacagta ttatgtagtc tgttttttat gcaaaatcta atttaatata ttgatattta   41760

tatcatttta cgtttctcgt tcagcttttct tgtacaaagt ggtgttaacc tagacttgtc   41820

catcttctgg attggccaac ttaattaatg tatgaaataa aaggatgcac acatagtgac   41880

atgctaatca ctataatgtg ggcatcaaag ttgtgtgtta tgtgtaatta ctagttatct   41940

gaataaaaga gaaagagatc atccatattt cttatcctaa atgaatgtca cgtgtcttta   42000

taattctttg atgaaccaga tgcatttcat taaccaaatc catatacata taaatattaa   42060

tcatatataa ttaatatcaa ttgggttagc aaaacaaatc tagtctaggt gtgttttgcg   42120

aattgcggcc gccaccgcgg tggagctcga attccggtcc gggtcacctt tgtccaccaa   42180

gatggaactg cggccgctca ttaattaagt caggcgcgcc tctagttgaa gacacgttca   42240

tgtcttcatc gtaagaagac actcagtagt cttcggccag aatggccatc tggattcagc   42300

aggcctagaa ggccatttaa atcctgagga tctggtcttc ctaaggaccc gggatatcgg   42360

accgattaaa ctttaattcg gtccgaagct tgaagttcct attccgaagt tcctattctc   42420

cagaaagtat aggaacttcg catgcctgca gtgcagcgtg acccggtcgt gcccctctct   42480

agagataatg agcattgcat gtctaagtta taaaaaatta ccacatattt tttttgtcac   42540

acttgtttga agtgcagttt atctatcttt atacatatat ttaaacttta ctctacgaat   42600

aatataatct atagtactac aataatatca gtgttttaga gaatcatata aatgaacagt   42660

tagacatggt ctaaaggaca attgagtatt ttgacaacag gactctacag ttttatcttt   42720

ttagtgtgca tgtgttctcc ttttttttttg caaatagctt cacctatata atacttcatc   42780

cattttatta gtacatccat ttagggttta gggttaatgg tttttataga ctaatttttt   42840

tagtacatct attttattct attttagcct ctaaattaag aaaactaaaa ctctatttta   42900

gtttttttat ttaataattt agatataaaa tagaataaaa taaagtgact aaaaattaaa   42960

caaataccct ttaagaaatt aaaaaaacta aggaaacatt tttcttgttt cgagtagata   43020

atgccagcct gttaaacgcc gtcgacgagt ctaacggaca ccaaccagcg aaccagcagc   43080

gtcgcgtcgg gccaagcgaa gcagacggca cggcatctct gtcgctgcct ctggacccct   43140

ctcgagagtt ccgctccacc gttggacttg ctccgctgtc ggcatccaga aattgcgtgg   43200

cggagcggca gacgtgagcc ggcacggcag gcggcctcct cctcctctca cggcaccggc   43260

agctacgggg gattcctttc ccaccgctcc ttcgctttcc cttcctcgcc cgccgtaata   43320

aatagacacc ccctccacac cctctttccc caacctcgtg ttgttcggag cgcacacaca   43380

cacaaccaga tctcccccaa atccaccgt cggcacctcc gcttcaaggt acgccgctcg   43440

tcctcccccc ccccctctc taccttctct agatcggcgt tccggtccat gcatggttag   43500

ggcccggtag ttctacttct gttcatgttt gtgttagatc cgtgtttgtg ttagatccgt   43560

gctgctagcg ttcgtacacg gatgcgacct gtacgtcaga cacgttctga ttgctaactt   43620
```

-continued

```
gccagtgttt ctctttgggg aatcctggga tggctctagc cgttccgcag acgggatcga   43680

tttcatgatt ttttttgttt cgttgcatag ggtttggttt gcccttttcc tttatttcaa   43740

tatatgccgt gcacttgttt gtcgggtcat cttttcatgc ttttttttgt cttggttgtg   43800

atgatgtggt ctggttgggc ggtcgttcta gatcggagta gaattctgtt tcaaactacc   43860

tggtggattt attaattttg gatctgtatg tgtgtgccat acatattcat agttacgaat   43920

tgaagatgat ggatggaaat atcgatctag gataggtata catgttgatg cgggtttttac  43980

tgatgcatat acagagatgc ttttttgttcg cttggttgtg atgatgtggt gtggttgggc   44040

ggtcgttcat tcgttctaga tcggagtaga atactgtttc aaactacctg gtgtatttat   44100

taattttgga actgtatgtg tgtgtcatac atcttcatag ttacgagttt aagatggatg   44160

gaaatatcga tctaggatag gtatacatgt tgatgtgggt tttactgatg catatacatg   44220

atggcatatg cagcatctat tcatatgctc taaccttgag tacctatcta ttataataaa   44280

caagtatgtt ttataattat tttgatcttg atatacttgg atgatggcat atgcagcagc   44340

tatatgtgga ttttttttagc cctgccttca tacgctattt atttgcttgg tactgtttct   44400

tttgtcgatg ctcaccctgt tgtttggtgt tacttctgca ggtcgacttt aacttagcct   44460

aggatccaca cgacaccatg atagaggtga aaccgattaa cgcagaggat acctatgaac   44520

taaggcatag aatactcaga ccaaaccagc cgatagaagc gtgtatgttt gaaagcgatt   44580

tacttcgtgg tgcatttcac ttaggcggct attacggggg caaactgatt tccatagctt   44640

cattccacca ggccgagcac tcagaactcc aaggccagaa acagtaccag ctccgaggta   44700

tggctacctt ggaaggttat cgtgagcaga aggcgggatc gagtctaatt aaacacgctg   44760

aagaaattct tcgtaagagg ggggcggact tgctttggtg taatgcgcgg acatccgcct   44820

caggctacta caaaaagtta ggcttcagcg agcagggaga ggtattcgac acgccgccag   44880

taggacctca catcctgatg tataaaagga tcacataact agctagtcag ttaacctaga   44940

cttgtccatc ttctggattg gccaacttaa ttaatgtatg aaataaaagg atgcacacat   45000

agtgacatgc taatcactat aatgtgggca tcaaagttgt gtgttatgtg taattactag   45060

ttatctgaat aaaagagaaa gagatcatcc atatttctta tcctaaatga atgtcacgtg   45120

tctttataat tctttgatga accagatgca tttcattaac caaatccata tacatataaa   45180

tattaatcat atataattaa tatcaattgg gttagcaaaa caaatctagt ctaggtgtgt   45240

tttgcgaatt cagagctcga attcattccg attaatcgtg gcctcttgct cttcaggatg   45300

aagagctatg tttaaacgtg caagcgctac tagacaattc agtacattaa aaacgtccgc   45360

aatgtgttat taagttgtct aagcgtcaat ttgtttacac cacaatatat cctgccacca   45420

gccagccaac agctccccga ccggcagctc ggcacaaaat caccactcga tacaggcagc   45480

ccatcagtcc gggacggcgt cagcgggaga gccgttgtaa ggcggcagac tttgctcatg   45540

ttaccgatgc tattcggaag aacggcaact aagctgccgg gtttgaaaca cggatgatct   45600

cgcggagggt agcatgttga ttgtaacgat gacagagcgt tgctgcctgt gatcaaatat   45660

catctccctc gcagagatcc gaattatcag ccttcttatt catttctcgc ttaaccgtga   45720

caggctgtcg atcttgagaa ctatgccgac ataataggaa atcgctggat aaagccgctg   45780

aggaagctga gtggcgctat ttctttagaa gtgaacgttg acgatcgtcg accgtacccc   45840

gatgaattaa ttcggacgta cgttctgaac acagctggat acttacttgg gcgattgtca   45900
```

-continued

```
tacatgacat caacaatgta cccgtttgtg taaccgtctc ttggaggttc gtatgacact   45960

agtggttccc ctcagcttgc gactagatgt tgaggcctaa cattttatta gagagcaggc   46020

tagttgctta gatacatgat cttcaggccg ttatctgtca gggcaagcga aaattggcca   46080

tttatgacga ccaatgcccc gcagaagctc ccatctttgc cgccatagac gccgcgcccc   46140

ccttttgggg tgtagaacat ccttttgcca gatgtggaaa agaagttcgt tgtcccattg   46200

ttggcaatga cgtagtagcc ggcgaaagtg cgagacccat ttgcgctata tataagccta   46260

cgatttccgt tgcgactatt gtcgtaattg gatgaactat tatcgtagtt gctctcagag   46320

ttgtcgtaat ttgatggact attgtcgtaa ttgcttatgg agttgtcgta gttgcttgga   46380

gaaatgtcgt agttggatgg ggagtagtca tagggaagac gagcttcatc cactaaaaca   46440

attggcaggt cagcaagtgc ctgccccgat gccatcgcaa gtacgaggct tagaaccacc   46500

ttcaacagat cgcgcatagt cttccccagc tctctaacgc ttgagttaag ccgcgccgcg   46560

aagcggcgtc ggcttgaacg aattgttaga cattatttgc cgactacctt ggtgatctcg   46620

cctttcacgt agtgaacaaa ttcttccaac tgatctgcgc gcgaggccaa gcgatcttct   46680

tgtccaagat aagcctgcct agcttcaagt atgacgggct gatactgggc cggcaggcgc   46740

tccattgccc agtcggcagc gacatccttc ggcgcgattt tgccggttac tgcgctgtac   46800

caaatgcggg acaacgtaag cactacattt cgctcatcgc cagcccagtc gggcggcgag   46860

ttccatagcg ttaaggtttc atttagcgcc tcaaatagat cctgttcagg aaccggatca   46920

aagagttcct ccgccgctgg acctaccaag gcaacgctat gttctcttgc ttttgtcagc   46980

aagatagcca gatcaatgtc gatcgtggct ggctcgaaga tacctgcaag aatgtcattg   47040

cgctgccatt ctccaaattg cagttcgcgc ttagctggat aacgccacgg aatgatgtcg   47100

tcgtgcacaa caatggtgac ttctacagcg cggagaatct cgctctctcc aggggaagcc   47160

gaagtttcca aaaggtcgtt gatcaaagct cgccgcgttg tttcatcaag ccttacagtc   47220

accgtaacca gcaaatcaat atcactgtgt ggcttcaggc cgccatccac tgcggagccg   47280

tacaaatgta cggccagcaa cgtcggttcg agatggcgct cgatgacgcc aactacctct   47340

gatagttgag tcgatacttc ggcgatcacc gcttccctca tgatgtttaa ctcctgaatt   47400

aagccgcgcc gcgaagcggt gtcggcttga atgaattgtt aggcgtcatc ctgtgctccc   47460

gagaaccagt accagtacat cgctgtttcg ttcgagactt gaggtctagt tttatacgtg   47520

aacaggtcaa tgccgccgag agtaaagcca cattttgcgt acaaattgca ggcaggtaca   47580

ttgttcgttt gtgtctctaa tcgtatgcca aggagctgtc tgcttagtgc ccactttttc   47640

gcaaattcga tgagactgtg cgcgactcct ttgcctcggt gcgtgtgcga cacaacaatg   47700

tgttcgatag aggctagatc gttccatgtt gagttgagtt caatcttccc gacaagctct   47760

tggtcgatga atgcgccata gcaagcagag tcttcatcag agtcatcatc cgagatgtaa   47820

tccttccggt aggggctcac acttctggta gatagttcaa agccttggtc ggataggtgc   47880

acatcgaaca cttcacgaac aatgaaatgg ttctcagcat ccaatgtttc cgccacctgc   47940

tcagggatca ccgaaatctt catatgacgc ctaacgcctg gcacagcgga tcgcaaacct   48000

ggcgcggctt ttggcacaaa aggcgtgaca ggtttgcgaa tccgttgctg ccacttgtta   48060

accctttttgc cagatttggt aactataatt tatgttagag gcgaagtctt gggtaaaaac   48120

tggcctaaaa ttgctgggga tttcaggaaa gtaaacatca ccttccggct cgatgtctat   48180
```

-continued

```
tgtagatata tgtagtgtat ctacttgatc gggggatctg ctgcctcgcg cgtttcggtg    48240

atgacggtga aaacctctga cacatgcagc tcccggagac ggtcacagct tgtctgtaag    48300

cggatgccgg gagcagacaa gcccgtcagg gcgcgtcagc gggtgttggc gggtgtcggg    48360

gcgcagccat gacccagtca cgtagcgata gcggagtgta tactggctta actatgcggc    48420

atcagagcag attgtactga gagtgcacca tatgcggtgt gaaataccgc acagatgcgt    48480

aaggagaaaa taccgcatca ggcgctcttc cgcttcctcg ctcactgact cgctgcgctc    48540

ggtcgttcgg ctgcggcgag cggtatcagc tcactcaaag gcggtaatac ggttatccac    48600

agaatcaggg gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa    48660

ccgtaaaaag gccgcgttgc tggcgttttt ccataggctc cgcccccctg acgagcatca    48720

caaaatcga cgctcaagtc agaggtggcg aaacccgaca ggactataaa gataccaggc    48780

gtttccccct ggaagctccc tcgtgcgctc tcctgttccg accctgccgc ttaccggata    48840

cctgtccgcc tttctccctt cgggaagcgt ggcgctttct catagctcac gctgtaggta    48900

tctcagttcg gtgtaggtcg ttcgctccaa gctgggctgt gtgcacgaac cccccgttca    48960

gcccgaccgc tgcgccttat ccggtaacta tcgtcttgag tccaacccgg taagacacga    49020

cttatcgcca ctggcagcag ccactggtaa caggattagc agagcgaggt atgtaggcgg    49080

tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagga cagtatttgg    49140

tatctgcgct ctgctgaagc cagttacctt cggaaaaaga gttggtagct cttgatccgg    49200

caaacaaacc accgctggta gcggtggttt ttttgtttgc aagcagcaga ttacgcgcag    49260

aaaaaaagga tctcaagaag atcctttgat cttttctacg gggtctgacg ctcagtggaa    49320

cgaaaactca cgttaaggga ttttggtcat gagattatca aaaaggatct tcacctagat    49380

ccttttaaat taaaaatgaa gttttaaatc aatctaaagt atatatgagt aaacttggtc    49440

tgacagttac caatgcttaa tcagtgaggc acctatctca gcgatctgtc tatttcgttc    49500

atccatagtt gcctgactcc ccgtcgtgta gataactacg atacgggagg gcttaccatc    49560

tggccccagt gctgcaatga taccgcgaga cccacgctca ccggctccag atttatcagc    49620

aataaaccag ccagccggaa gggccgagcg cagaagtggt cctgcaactt tatccgcctc    49680

catccagtct attaattgtt gccgggaagc tagagtaagt agttcgccag ttaatagttt    49740

gcgcaacgtt gttgccattg ctgca                                         49765
```

```
<210> SEQ ID NO 44
<211> LENGTH: 12856
<212> TYPE: DNA
<213> ORGANISM: artificial
<220> FEATURE:
<223> OTHER INFORMATION: vector

<400> SEQUENCE: 44
```

```
cgccttggcg cgccgatcat ccacaagttt gtacaaaaaa gctgaacgag aaacgtaaaa     60

tgatataaat atcaatatat taaattagat tttgcataaa aaacagacta cataatactg    120

taaaacacaa catatccagt cactatggcg gccgcattag caccccagg ctttacactt    180

tatgcttccg gctcgtataa tgtgtggatt ttgagttagg atttaaatac gcgttgatcc    240

ggcttactaa aagccagata acagtatgcg tatttgcgcg ctgatttttg cggtataaga    300

atatatactg atatgtatac ccgaagtatg tcaaaaagag gtatgctatg aagcagcgta    360
```

-continued

```
ttacagtgac agttgacagc gacagctatc agttgctcaa ggcatatatg atgtcaatat   420

ctccggtctg gtaagcacaa ccatgcagaa tgaagcccgt cgtctgcgtg ccgaacgctg   480

gaaagcggaa aatcaggaag ggatggctga ggtcgcccgg tttattgaaa tgaacggctc   540

ttttgctgac gagaacaggg gctggtgaaa tgcagtttaa ggtttacacc tataaaagag   600

agagccgtta tcgtctgttt gtggatgtac agagtgatat cattgacacg cccggtcgac   660

ggatggtgat cccccctggcc agtgcacgtc tgctgtcaga taaagtctcc cgtgaacttt   720

acccggtggt gcatatcggg gatgaaagct ggcgcatgat gaccaccgat atggccagtg   780

tgccggtctc cgttatcggg gaagaagtgg ctgatctcag ccaccgcgaa aatgacatca   840

aaaacgccat taacctgatg ttctggggaa tataaatgtc aggctcccctt atacacagcc   900

agtctgcagg tcgaccatag tgactggata tgttgtgttt tacagtatta tgtagtctgt   960

tttttatgca aaatctaatt taatatattg atatttatat cattttacgt ttctcgttca  1020

gctttcttgt acaaagtggt gttaacctag acttgtccat cttctggatt ggccaactta  1080

attaatgtat gaaataaaag gatgcacaca tagtgacatg ctaatcacta taatgtgggc  1140

atcaaagttg tgtgttatgt gtaattacta gttatctgaa taaaagagaa agagatcatc  1200

catatttctt atcctaaatg aatgtcacgt gtctttataa ttctttgatg aaccagatgc  1260

atttcattaa ccaaatccat atacatataa atattaatca tatataatta atatcaattg  1320

ggttagcaaa acaaatctag tctaggtgtg ttttgcgaat tgcggccgcc accgcggtgg  1380

agctcgaatt ccggtccggg tcacctttgt ccaccaagat ggaactgcgg ccgctcatta  1440

attaagtcag gcgcgcctct agttgaagac acgttcatgt cttcatcgta agaagacact  1500

cagtagtctt cggccagaat ggccatctgg attcagcagg cctagaaggc catttaaatc  1560

ctgaggatct ggtcttccta aggacccggg atatcggacc gattaaactt taattcggtc  1620

cgaagcttga agttcctatt ccgaagttcc tattctccag aaagtatagg aacttcgcat  1680

gcctgcagtg cagcgtgacc cggtcgtgcc cctctctaga gataatgagc attgcatgtc  1740

taagttataa aaaattacca catatttttt ttgtcacact tgtttgaagt gcagtttatc  1800

tatctttata catatattta aactttactc tacgaataat ataatctata gtactacaat  1860

aatatcagtg ttttagagaa tcatataaat gaacagttag acatggtcta aaggacaatt  1920

gagtattttg acaacaggac tctacagttt tatctttta gtgtgcatgt gttctccttt  1980

ttttttgcaa atagcttcac ctatataata cttcatccat tttattagta catccattta  2040

gggtttaggg ttaatggttt ttatagacta atttttttag tacatctatt ttattctatt  2100

ttagcctcta aattaagaaa actaaaactc tattttagtt tttttattta ataatttaga  2160

tataaaatag aataaaataa agtgactaaa aattaaacaa ataccctttaa agaaattaaa  2220

aaaactaagg aaacattttt cttgtttcga gtagataatg ccagcctgtt aaacgccgtc  2280

gacgagtcta acggacacca accagcgaac cagcagcgtc gcgtcgggcc aagcgaagca  2340

gacggcacgg catctctgtc gctgcctctg gacccctctc gagagttccg ctccaccgtt  2400

ggacttgctc cgctgtcggc atccagaaat tgcgtggcgg agcggcagac gtgagccggc  2460

acggcaggcg gcctcctcct cctctcacgg caccggcagc tacggggat tcctttccca  2520

ccgctccttc gctttccctt cctcgcccgc cgtaataaat agacaccccc tccacaccct  2580

ctttccccaa cctcgtgttg ttcggagcgc acacacacac aaccagatct ccccaaatc  2640
```

-continued

```
caccgtcgg cacctccgct tcaaggtacg ccgctcgtcc tcccccccc ccctctctac      2700

cttctctaga tcggcgttcc ggtccatgca tggttagggc ccggtagttc tacttctgtt      2760

catgtttgtg ttagatccgt gtttgtgtta gatccgtgct gctagcgttc gtacacggat      2820

gcgacctgta cgtcagacac gttctgattg ctaacttgcc agtgtttctc tttgggaat      2880

cctgggatgg ctctagccgt tccgcagacg ggatcgattt catgattttt tttgtttcgt      2940

tgcatagggt ttggtttgcc cttttccttt atttcaatat atgccgtgca cttgtttgtc      3000

gggtcatctt ttcatgcttt tttttgtctt ggttgtgatg atgtggtctg gttgggcggt      3060

cgttctagat cggagtagaa ttctgtttca aactacctgg tggatttatt aattttggat      3120

ctgtatgtgt gtgccataca tattcatagt tacgaattga agatgatgga tggaaatatc      3180

gatctaggat aggtatacat gttgatgcgg gttttactga tgcatataca gagatgcttt      3240

ttgttcgctt ggttgtgatg atgtggtgtg gttgggcggt cgttcattcg ttctagatcg      3300

gagtagaata ctgtttcaaa ctacctggtg tatttattaa ttttggaact gtatgtgtgt      3360

gtcatacatc ttcatagtta cgagtttaag atggatggaa atatcgatct aggataggta      3420

tacatgttga tgtgggtttt actgatgcat atacatgatg gcatatgcag catctattca      3480

tatgctctaa ccttgagtac ctatctatta taataaacaa gtatgtttta taattatttt      3540

gatcttgata tacttggatg atggcatatg cagcagctat atgtggattt ttttagccct      3600

gccttcatac gctatttatt tgcttggtac tgtttctttt gtcgatgctc accctgttgt      3660

ttggtgttac ttctgcaggt cgactttaac ttagcctagg atccacacga caccatgata      3720

gaggtgaaac cgattaacgc agaggatacc tatgaactaa ggcatagaat actcagacca      3780

aaccagccga tagaagcgtg tatgtttgaa agcgatttac ttcgtggtgc atttcactta      3840

ggcggctatt acggggcaa actgatttcc atagcttcat tccaccaggc cgagcactca      3900

gaactccaag gccagaaaca gtaccagctc cgaggtatgg ctaccttgga aggttatcgt      3960

gagcagaagg cgggatcgag tctaattaaa cacgctgaag aaattcttcg taagaggggg      4020

gcggacttgc tttggtgtaa tgcgcggaca tccgcctcag gctactacaa aaagttaggc      4080

ttcagcgagc agggagaggt attcgacacg ccgccagtag gacctcacat cctgatgtat      4140

aaaaggatca cataactagc tagtcagtta acctagactt gtccatcttc tggattggcc      4200

aacttaatta atgtatgaaa taaaaggatg cacacatagt gacatgctaa tcactataat      4260

gtgggcatca aagttgtgtg ttatgtgtaa ttactagtta tctgaataaa agagaaagag      4320

atcatccata tttcttatcc taaatgaatg tcacgtgtct ttataattct ttgatgaacc      4380

agatgcattt cattaaccaa atccatatac atataaatat taatcatata taattaatat      4440

caattgggtt agcaaaacaa atctagtcta ggtgtgtttt gcgaattcag agctcgaatt      4500

cattccgatt aatcgtggcc tcttgctctt caggatgaag agctatgttt aaacgtgcaa      4560

gcgctactag acaattcagt acattaaaaa cgtccgcaat gtgttattaa gttgtctaag      4620

cgtcaatttg tttacaccac aatatatcct gccaccagcc agccaacagc tccccgaccg      4680

gcagctcggc acaaaatcac cactcgatac aggcagccca tcagtccggg acggcgtcag      4740

cgggagagcc gttgtaaggc ggcagacttt gctcatgtta ccgatgctat tcggaagaac      4800

ggcaactaag ctgccgggtt tgaaacacgg atgatctcgc ggagggtagc atgttgattg      4860

taacgatgac agagcgttgc tgcctgtgat caaatatcat ctccctcgca gagatccgaa      4920
```

-continued

```
ttatcagcct tcttattcat ttctcgctta accgtgacag gctgtcgatc ttgagaacta     4980

tgccgacata ataggaaatc gctggataaa gccgctgagg aagctgagtg gcgctatttc     5040

tttagaagtg aacgttgacg atcgtcgacc gtaccccgat gaattaattc ggacgtacgt     5100

tctgaacaca gctggatact tacttgggcg attgtcatac atgacatcaa caatgtaccc     5160

gtttgtgtaa ccgtctcttg gaggttcgta tgacactagt ggttcccctc agcttgcgac     5220

tagatgttga ggcctaacat tttattagag agcaggctag ttgcttagat acatgatctt     5280

caggccgtta tctgtcaggg caagcgaaaa ttggccattt atgacgacca atgccccgca     5340

gaagctccca tctttgccgc catagacgcc gcgcccccct tttggggtgt agaacatcct     5400

tttgccagat gtggaaaaga agttcgttgt cccattgttg gcaatgacgt agtagccggc     5460

gaaagtgcga gacccatttg cgctatatat aagcctacga tttccgttgc gactattgtc     5520

gtaattggat gaactattat cgtagttgct ctcagagttg tcgtaatttg atggactatt     5580

gtcgtaattg cttatggagt tgtcgtagtt gcttggagaa atgtcgtagt tggatgggga     5640

gtagtcatag ggaagacgag cttcatccac taaaacaatt ggcaggtcag caagtgcctg     5700

ccccgatgcc atcgcaagta cgaggcttag aaccaccttc aacagatcgc gcatagtctt     5760

ccccagctct ctaacgcttg agttaagccg cgccgcgaag cggcgtcggc ttgaacgaat     5820

tgttagacat tatttgccga ctaccttggt gatctcgcct ttcacgtagt gaacaaattc     5880

ttccaactga tctgcgcgcg aggccaagcg atcttcttgt ccaagataag cctgcctagc     5940

ttcaagtatg acgggctgat actgggccgg caggcgctcc attgcccagt cggcagcgac     6000

atccttcggc gcgattttgc cggttactgc gctgtaccaa atgcgggaca acgtaagcac     6060

tacatttcgc tcatcgccag cccagtcggg cggcgagttc catagcgtta aggtttcatt     6120

tagcgcctca aatagatcct gttcaggaac cggatcaaag agttcctccg ccgctggacc     6180

taccaaggca acgctatgtt ctcttgcttt tgtcagcaag atagccagat caatgtcgat     6240

cgtggctggc tcgaagatac ctgcaagaat gtcattgcgc tgccattctc caaattgcag     6300

ttcgcgctta gctggataac gccacggaat gatgtcgtcg tgcacaacaa tggtgacttc     6360

tacagcgcgg agaatctcgc tctctccagg ggaagccgaa gtttccaaaa ggtcgttgat     6420

caaagctcgc cgcgttgttt catcaagcct tacagtcacc gtaaccagca aatcaatatc     6480

actgtgtggc ttcaggccgc catccactgc ggagccgtac aaatgtacgg ccagcaacgt     6540

cggttcgaga tggcgctcga tgacgccaac tacctctgat agttgagtcg atacttcggc     6600

gatcaccgct tccctcatga tgtttaactc ctgaattaag ccgcgccgcg aagcggtgtc     6660

ggcttgaatg aattgttagg cgtcatcctg tgctcccgag aaccagtacc agtacatcgc     6720

tgtttcgttc gagacttgag gtctagtttt atacgtgaac aggtcaatgc cgccgagagt     6780

aaagccacat tttgcgtaca aattgcaggc aggtacattg ttcgtttgtg tctctaatcg     6840

tatgccaagg agctgtctgc ttagtgccca ctttttcgca aattcgatga gactgtgcgc     6900

gactcctttg cctcggtgcg tgtgcgacac aacaatgtgt cgatagagg ctagatcgtt     6960

ccatgttgag ttgagttcaa tcttcccgac aagctcttgg tcgatgaatg cgccatagca     7020

agcagagtct tcatcagagt catcatccga gatgtaatcc ttccggtagg ggctcacact     7080

tctggtagat agttcaaagc cttggtcgga taggtgcaca tcgaacactt cacgaacaat     7140

gaaatggttc tcagcatcca atgtttccgc cacctgctca gggatcaccg aaatcttcat     7200
```

-continued

```
atgacgccta acgcctggca cagcggatcg caaacctggc gcggcttttg gcacaaaagg    7260

cgtgacaggt ttgcgaatcc gttgctgcca cttgttaacc cttttgccag atttggtaac    7320

tataatttat gttagaggcg aagtcttggg taaaaactgg cctaaaattg ctggggattt    7380

caggaaagta aacatcacct tccggctcga tgtctattgt agatatatgt agtgtatcta    7440

cttgatcggg ggatctgctg cctcgcgcgt ttcggtgatg acggtgaaaa cctctgacac    7500

atgcagctcc cggagacggt cacagcttgt ctgtaagcgg atgccgggag cagacaagcc    7560

cgtcagggcg cgtcagcggg tgttggcggg tgtcggggcg cagccatgac ccagtcacgt    7620

agcgatagcg gagtgtatac tggcttaact atgcggcatc agagcagatt gtactgagag    7680

tgcaccatat gcggtgtgaa ataccgcaca gatgcgtaag gagaaaatac cgcatcaggc    7740

gctcttccgc ttcctcgctc actgactcgc tgcgctcggt cgttcggctg cggcgagcgg    7800

tatcagctca ctcaaaggcg gtaatacggt tatccacaga atcaggggat aacgcaggaa    7860

agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg    7920

cgtttttcca taggctccgc cccctgacg agcatcacaa aaatcgacgc tcaagtcaga    7980

ggtggcgaaa cccgacagga ctataaagat accaggcgtt tccccctgga agctccctcg    8040

tgcgctctcc tgttccgacc ctgccgctta ccggatacct gtccgccttt ctcccttcgg    8100

gaagcgtggc gctttctcat agctcacgct gtaggtatct cagttcggtg taggtcgttc    8160

gctccaagct gggctgtgtg cacgaacccc cgttcagcc cgaccgctgc gccttatccg    8220

gtaactatcg tcttgagtcc aacccggtaa gacacgactt atcgccactg gcagcagcca    8280

ctggtaacag gattagcaga gcgaggtatg taggcggtgc tacagagttc ttgaagtggt    8340

ggcctaacta cggctacact agaaggacag tatttggtat ctgcgctctg ctgaagccag    8400

ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa acaaaccacc gctggtagcg    8460

gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa aaaaggatct caagaagatc    8520

ctttgatctt ttctacgggg tctgacgctc agtggaacga aaactcacgt taagggattt    8580

tggtcatgag attatcaaaa aggatcttca cctagatcct tttaaattaa aaatgaagtt    8640

ttaaatcaat ctaaagtata tatgagtaaa cttggtctga cagttaccaa tgcttaatca    8700

gtgaggcacc tatctcagcg atctgtctat ttcgttcatc catagttgcc tgactccccg    8760

tcgtgtagat aactacgata cgggagggct taccatctgg ccccagtgct gcaatgatac    8820

cgcgagaccc acgctcaccg gctccagatt tatcagcaat aaaccagcca gccggaaggg    8880

ccgagcgcag aagtggtcct gcaactttat ccgcctccat ccagtctatt aattgttgcc    8940

gggaagctag agtaagtagt tcgccagtta atagtttgcg caacgttgtt gccattgctg    9000

caggggggg gggggggggg gacttccatt gttcattcca cggacaaaaa cagagaaagg    9060

aaacgacaga ggccaaaaag cctcgctttc agcacctgtc gtttcctttc ttttcagagg    9120

gtattttaaa taaaaacatt aagttatgac gaagaagaac ggaaacgcct aaaccggaa    9180

aattttcata aatagcgaaa acccgcgagg tcgccgcccc gtaacctgtc ggatcaccgg    9240

aaaggacccg taaagtgata atgattatca tctacatatc acaacgtgcg tggaggccat    9300

caaaccacgt caaataatca attatgacgc aggtatcgta ttaattgatc tgcatcaact    9360

taacgtaaaa acaacttcag acaatacaaa tcagcgacac tgaatacggg gcaacctcat    9420

gtccccccc cccccccccc tgcaggcatc gtggtgtcac gctcgtcgtt tggtatggct    9480
```

-continued

```
tcattcagct ccggttccca acgatcaagg cgagttacat gatcccccat gttgtgcaaa      9540

aaagcggtta gctccttcgg tcctccgatc gttgtcagaa gtaagttggc cgcagtgtta      9600

tcactcatgg ttatggcagc actgcataat tctcttactg tcatgccatc cgtaagatgc      9660

ttttctgtga ctggtgagta ctcaaccaag tcattctgag aatagtgtat gcggcgaccg      9720

agttgctctt gcccggcgtc aacacgggat aataccgcgc cacatagcag aactttaaaa      9780

gtgctcatca ttggaaaacg ttcttcgggg cgaaaactct caaggatctt accgctgttg      9840

agatccagtt cgatgtaacc cactcgtgca cccaactgat cttcagcatc ttttactttc      9900

accagcgttt ctgggtgagc aaaaacagga aggcaaaatg ccgcaaaaaa gggaataagg      9960

gcgacacgga aatgttgaat actcatactc ttcctttttc aatattattg aagcatttat     10020

cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaaata     10080

ggggttccgc gcacatttcc ccgaaaagtg ccacctgacg tctaagaaac cattattatc     10140

atgacattaa cctataaaaa taggcgtatc acgaggccct ttcgtcttca agaattggtc     10200

gacgatcttg ctgcgttcgg atattttcgt ggagttcccg ccacagaccc ggattgaagg     10260

cgagatccag caactcgcgc cagatcatcc tgtgacggaa ctttggcgcg tgatgactgg     10320

ccaggacgtc ggccgaaaga gcgacaagca gatcacgctt ttcgacagcg tcggatttgc     10380

gatcgaggat ttttcggcgc tgcgctacgt ccgcgaccgc gttgagggat caagccacag     10440

cagcccactc gaccttctag ccgacccaga cgagccaagg gatctttttg gaatgctgct     10500

ccgtcgtcag gctttccgac gtttgggtgg ttgaacagaa gtcattatcg tacggaatgc     10560

caagcactcc cgaggggaac cctgtggttg gcatgcacat acaaatggac gaacggataa     10620

acctttccac gccctttaa atatccgtta ttctaataaa cgctctttc tcttaggttt      10680

acccgccaat atatcctgtc aaacactgat agtttaaact gaaggcggga aacgacaatc     10740

tgatcatgag cggagaatta agggagtcac gttatgaccc ccgccgatga cgcgggacaa     10800

gccgttttac gtttggaact gacagaaccg caacgttgaa ggagccactc agcaagctgg     10860

tacgattgta atacgactca ctatagggcg aattgagcgc tgtttaaacg ctcttcaact     10920

ggaagagcgg ttacccggac cgaagcttga agttcctatt ccgaagttcc tattctctag     10980

aaagtatagg aacttcagat ctcgatgctc accctgttgt ttggtgttac ttctgcaggt     11040

cgactctaga ggatccacca tgagcccaga acgacgcccg gccgcatcc gccgtgccac      11100

cgaggcggac atgccggcgg tctgcaccat cgtcaaccac tacatcgaga caagcacggt     11160

caacttccgt accgagccgc aggaaccgca ggactggacg gacgacctcg tccgtctgcg     11220

ggagcgctat ccctggctcg tcgccgaggt ggacggcgag gtcgccggca tcgcctacgc     11280

gggcccctgg aaggcacgca acgcctacga ctggacggcc gagtcgaccg tgtacgtctc     11340

cccccgccac cagcggacgg gactgggctc cacgctctac acccacctgc tgaagtccct     11400

ggaggcacag ggcttcaaga gcgtggtcgc tgtcatcggg ctgcccaacg acccgagcgt     11460

gcgcatgcac gaggcgctcg gatatgcccc ccgcggcatg ctgcgggcgg ccggcttcaa     11520

gcacgggaac tggcatgacg tgggtttctg gcagctggac ttcagcctgc cggtaccgcc     11580

ccgtccggtc ctgcccgtca ccgagatctg atccgtcgac caacctagac ttgtccatct     11640

tctggattgg ccaacttaat taatgtatga aataaaagga tgcacacata gtgacatgct     11700

aatcactata atgtgggcat caaagttgtg tgttatgtgt aattactagt tatctgaata     11760
```

-continued

```
aaagagaaag agatcatcca tatttcttat cctaaatgaa tgtcacgtgt ctttataatt    11820

cttttgatgaa ccagatgcat ttcattaacc aaatccatat acatataaat attaatcata   11880

tataattaat atcaattggg ttagcaaaac aaatctagtc taggtgtgtt ttgcgaattg    11940

cggccgcgat ctggggaatt cccatggaca ccggtaattc ccatgatctt ctctccttca    12000

tcaatggatg ccatgtttca taacaataac accaaatgtt tgatgagcta ccaacaattg    12060

cgcaaagact atggctaagc tcgagctcgc tcgctacaag ttgttgactt tcaaatacaa    12120

gtttgttttt ggaacaccaa atattctaca tgatctttca ctaagttgcg caccactatc    12180

aaaagattat ctaggccatt attcaagtaa agagtgaaca cgtctaagac ccacaaccac    12240

accaaataga atacgcatac atgcaacata ttgtgcaaga agtatccaac tggactccca    12300

tgtattctaa aactattttc gtagagttaa agttatgaca aacttatcaa ataaaaattt    12360

gaacgctgga ccaaaacttt catctttcaa atccaccatc gtctatcctc ataaattgtt    12420

ttgattataa cacatctacg taaatcattt gttttgaaca atactaattt aattttatta    12480

agtcaaataa cctgcttaga aaataatccc tccacctcat ttaacaattt cttgtcaaac    12540

acacaccaag aaaaaaatta atgaaagaga aagaaatga aaaggacatg gagttgaata     12600

ctagcaaaat tgattgaagg aagattcaca attgaaattg aaaccatta atttatttc       12660

gggtccataa taataaattg gtaagaataa aaacccgatc aagtccggta cagtacaatt    12720

ccactccacc aactccttac ttaaacccct atttatacc actctcatcc tcactcttcc       12780

ttcacctctc acactctctt ctctctctca aaaccctcac acaaacgctg cgtttagtgt    12840

aagaaattca atccgg                                                      12856
```

```
<210> SEQ ID NO 45
<211> LENGTH: 879
<212> TYPE: DNA
<213> ORGANISM: Medicago sativa

<400> SEQUENCE: 45
```

```
aattcccatg atcttctctc cttcatcaat ggatgccatg tttcataaca ataacaccaa       60

atgtttgatg agctaccaac aattgcgcaa agactatggc taagctcgag ctcgctcgct     120

acaagttgtt gactttcaaa tacaagtttg tttttggaac accaaatatt ctacatgatc    180

tttcactaag ttgcgcacca ctatcaaaag attatctagg ccattattca agtaaagagt     240

gaacacgtct aagacccaca accacaccaa atagaatacg catacatgca acatattgtg     300

caagaagtat ccaactggac tcccatgtat tctaaaacta ttttcgtaga gttaaagtta    360

tgacaaactt atcaaataaa aatttgaacg ctggaccaaa actttcatct ttcaaatcca    420

ccatcgtcta tcctcataaa ttgttttgat ataacacat ctacgtaaat catttgtttt      480

gaacaatact aatttaattt tattaagtca ataacctgc ttagaaaata atccctccac      540

ctcatttaac aatttcttgt caaacacaca ccagaaaaaa aattaatgaa agagaaaga     600

aatgaaaagg acatggagtt gaatactagc aaaattgatt gaaggaagat tcacaattga    660

aattgaaacc atttaattta ttttcgggtc cataataata aattggtaag aataaaaacc     720
```

-continued

```
cgatcaagtc cggtacagta caattccact ccaccaactc cttacttaaa cccctattta     780

tacccactct catcctcact cttccttcac ctctcacact ctcttctctc tctcaaaacc     840

ctcacacaaa cgctgcgttt agtgtaagaa attcaatcc                            879
```

What is claimed is:

1. A plant comprising in its genome a recombinant DNA construct comprising a polynucleotide operably linked to at least one regulatory element, wherein said polynucleotide encodes a polypeptide having an amino acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13, 17, 19, or 21 and wherein said plant exhibits altered root architecture when compared to a control plant not comprising said recombinant DNA construct.

2. The plant of claim 1, wherein the plant is a maize plant or a soybean plant.

3. A plant comprising in its genome:

a recombinant DNA construct comprising:

(a) a polynucleotide operably linked to at least one regulatory element, wherein said polynucleotide encodes a polypeptide having an amino acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13, 17, 19, or 21, or

(b) a suppression DNA construct comprising at least one regulatory element operably linked to:

(i) all or part of:

(A) a nucleic acid sequence encoding a polypeptide having an amino acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13, 17, 19, or 21, or

(B) a full complement of the nucleic acid sequence of (b)(i)(A); or

(ii) a region derived from all or part of a sense strand or antisense strand of a target gene of interest, said region having a nucleic acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to said all or part of a sense strand or antisense strand from which said region is derived, and wherein said target gene of interest encodes an RT1 or RT1-like polypeptide, and wherein said plant exhibits an alteration of at least one agronomic characteristic when compared to a control plant not comprising said recombinant DNA construct.

4. The plant of claim 3, wherein the plant is a maize plant or a soybean plant.

5. The plant of claim 3, wherein said plant exhibits said alteration of said at least one agronomic characteristic when compared, under varying environmental conditions, to said control plant not comprising said recombinant DNA construct.

6. The plant of claim 5, wherein said varying environmental condition is at least one selected from drought, nitrogen, soil type, insect or disease.

7. The plant of claim 5, wherein the plant is a maize plant or a soybean plant.

8. The plant of claim 6, wherein the plant is a maize plant or a soybean plant.

9. The plant of claim 7, wherein the plant is a maize plant or a soybean plant.

10. The plant of claim 3, wherein said at least one agronomic characteristic is selected from the group consisting of greenness, yield, growth rate, biomass, fresh weight at maturation, dry weight at maturation, fruit yield, seed yield, total plant nitrogen content, fruit nitrogen content, seed nitrogen content, nitrogen content in a vegetative tissue, total plant free amino acid content, fruit free amino acid content, seed free amino acid content, free amino acid content in a vegetative tissue, total plant protein content, fruit protein content, seed protein content, protein content in a vegetative tissue, drought tolerance, nitrogen uptake, root penetration, root lodging, stalk lodging, plant height, ear length, and harvest index.

11. The plant of claim 10, wherein the plant is a maize plant or a soybean plant.

12. The plant of claim 3, wherein said plant exhibits an increase of said at least one agronomic characteristic when compared to said control plant.

13. The plant of claim 12, wherein the plant is a maize plant or a soybean plant.

14. A method of altering root architecture in a plant, comprising:

(a) introducing into a regenerable plant cell a recombinant DNA construct comprising a polynucleotide operably linked to at least one regulatory sequence, wherein the polynucleotide encodes a polypeptide having an amino acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13, 17, 19, or 21; and

(b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the recombinant DNA construct and exhibits altered root architecture when compared to a control plant not comprising the recombinant DNA construct.

15. The method of claim 14, further comprising:

(c) obtaining a progeny plant derived from the transgenic plant, wherein said progeny plant comprises in its genome the recombinant DNA construct and exhibits altered root architecture when compared to a control plant not comprising the recombinant DNA construct.

16. A method of evaluating root architecture in a plant, comprising:

(a) introducing into a regenerable plant cell a recombinant DNA construct comprising a polynucleotide operably linked to at least one regulatory sequence, wherein the polynucleotide encodes a polypeptide having an amino acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13, 17, 19, or 21;

(b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the recombinant DNA construct; and

(c) evaluating root architecture of the transgenic plant compared to a control plant not comprising the recombinant DNA construct.

17. The method of claim 16, further comprising:

(d) obtaining a progeny plant derived from the transgenic plant, wherein the progeny plant comprises in its genome the recombinant DNA construct; and

(e) evaluating root architecture of the progeny plant compared to a control plant not comprising the recombinant DNA construct.

18. A method of evaluating root architecture in a plant, comprising:

(a) introducing into a regenerable plant cell a recombinant DNA construct comprising a polynucleotide operably linked to at least one regulatory sequence, wherein the polynucleotide encodes a polypeptide having an amino acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13, 17, 19, or 21;

(b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the recombinant DNA construct;

(c) obtaining a progeny plant derived from the transgenic plant, wherein the progeny plant comprises in its genome the recombinant DNA construct; and

(d) evaluating root architecture of the progeny plant compared to a control plant not comprising the recombinant DNA construct.

19. A method of determining an alteration of an agronomic characteristic in a plant, comprising:

(a) introducing into a regenerable plant cell a recombinant DNA construct comprising a polynucleotide operably linked to at least one regulatory sequence, wherein the polynucleotide encodes a polypeptide having an amino acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13, 17, 19, or 21;

(b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the recombinant DNA construct; and

(c) determining whether the transgenic plant exhibits an alteration of at least one agronomic characteristic when compared to a control plant not comprising the recombinant DNA construct.

20. The method of claim 19, further comprising:

(d) obtaining a progeny plant derived from the transgenic plant, wherein the progeny plant comprises in its genome the recombinant DNA construct; and

(e) determining whether the progeny plant exhibits an alteration of at least one agronomic characteristic when compared to a control plant not comprising the recombinant DNA construct.

21. The method of claim 19, wherein said determining step comprises determining whether the transgenic plant exhibits an alteration of at least one agronomic characteristic when compared, under varying environmental conditions, to a control plant not comprising the recombinant DNA construct.

22. The method of claim 20, wherein said varying environmental condition is at least one selected from drought, nitrogen, soil type, insect or disease.

23. The method of claim 20, wherein step (e) comprises determining whether the progeny plant exhibits an alteration of at least one agronomic characteristic when compared, under varying environmental conditions, to a control plant not comprising the recombinant DNA construct.

24. The method of claim 23, wherein said varying environmental condition is at least one selected from drought, nitrogen, soil type, insect or disease.

25. A method of determining an alteration of an agronomic characteristic in a plant, comprising:

(a) introducing into a regenerable plant cell a recombinant DNA construct comprising a polynucleotide operably linked to at least one regulatory sequence, wherein the polynucleotide encodes a polypeptide having an amino acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13, 17, 19, or 21;

(b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the recombinant DNA construct;

(c) obtaining a progeny plant derived from the transgenic plant, wherein the progeny plant comprises in its genome the recombinant DNA construct; and

(d) determining whether the progeny plant exhibits an alteration of at least one agronomic characteristic when compared to a control plant not comprising the recombinant DNA construct.

26. The method of claim 25, wherein step (d) comprises determining whether the transgenic plant exhibits an alteration of at least one agronomic characteristic when compared, under varying environmental conditions, to a control plant not comprising the recombinant DNA construct.

27. A method of determining an alteration of an agronomic characteristic in a plant, comprising:

(a) introducing into a regenerable plant cell a suppression DNA construct comprising at least one regulatory element operably linked to:

(i) all or part of:

(A) a nucleic acid sequence encoding a polypeptide having an amino acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13, 17, 19, or 21, or

(B) a full complement of the nucleic acid sequence of (a)(i)(A); or

(ii) a region derived from all or part of a sense strand or antisense strand of a target gene of interest, said region having a nucleic acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to said all or part of a sense strand or antisense strand from which said region is derived, and wherein said target gene of interest encodes a RT1 or RT1-like polypeptide;

(b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the suppression DNA construct; and

(c) determining whether the transgenic plant exhibits an alteration of at least one agronomic characteristic when compared to a control plant not comprising the suppression DNA construct.

28. The method of claim 27, wherein said determining step comprises determining whether the transgenic plant exhibits an alteration of at least one agronomic characteristic when compared, under varying environmental conditions, to a control plant not comprising the suppression DNA construct.

29. The method of claim 28, wherein said varying environmental condition is at least one selected from drought, nitrogen, soil type, insect or disease.

30. The method of claim 27, further comprising:

(d) obtaining a progeny plant derived from the transgenic plant, wherein the progeny plant comprises in its genome the suppression DNA construct; and

(e) determining whether the progeny plant exhibits an alteration of at least one agronomic characteristic when compared to a control plant not comprising the suppression DNA construct.

31. The method of claim 30, wherein step (e) comprises determining whether the progeny plant exhibits an alteration of at least one agronomic characteristic when compared, under varying environmental conditions, to a control plant not comprising the suppression DNA construct.

32. The method of claim 31, wherein said varying environmental condition is at least one selected from drought, nitrogen, soil type, insect or disease.

33. A method of determining an alteration of an agronomic characteristic in a plant, comprising:

(a) introducing into a regenerable plant cell a suppression DNA construct comprising at least one regulatory element operably linked to:

(i) all or part of:

(A) a nucleic acid sequence encoding a polypeptide having an amino acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13, 17, 19, or 21, or

(B) a full complement of the nucleic acid sequence of (a)(i)(A); or

(ii) a region derived from all or part of a sense strand or antisense strand of a target gene of interest, said region having a nucleic acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to said all or part of a sense strand or antisense strand from which said region is derived, and wherein said target gene of interest encodes a RT1 or RT1-like polypeptide;

(b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the suppression DNA construct and exhibits altered root architecture when compared to a control plant not comprising the suppression DNA construct;

(c) obtaining a progeny plant derived from the transgenic plant, wherein the progeny plant comprises in its genome the suppression DNA construct; and

(d) determining whether the progeny plant exhibits an alteration of at least one agronomic characteristic when compared to a control plant not comprising the suppression DNA construct.

34. The method of claim 33, wherein step (d) comprises determining whether the transgenic plant exhibits an alter-

ation of at least one agronomic characteristic when compared, under varying environmental conditions, to a control plant not comprising the recombinant DNA construct.

35. The method of claim 34, wherein said varying environmental condition is at least one selected from drought, nitrogen, soil type, insect or disease.

36. A method of altering root architecture in a plant, comprising:

(a) introducing into a regenerable plant cell a suppression DNA construct comprising at least one regulatory element operably linked to:

(i) all or part of:

(A) a nucleic acid sequence encoding a polypeptide having an amino acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13, 17, 19, or 21, or

(B) a full complement of the nucleic acid sequence of (a)(i)(A); or

(ii) a region derived from all or part of a sense strand or antisense strand of a target gene of interest, said region having a nucleic acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to said all or part of a sense strand or antisense strand from which said region is derived, and wherein said target gene of interest encodes a RT1 or RT1-like polypeptide; and

(b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the suppression DNA construct and wherein the transgenic plant exhibits altered root architecture when compared to a control plant not comprising the suppression DNA construct.

37. The method of claim 36, further comprising:

(c) obtaining a progeny plant derived from the transgenic plant, wherein said progeny plant comprises in its genome the recombinant DNA construct and wherein the progeny plant exhibits altered root architecture when compared to a control plant not comprising the suppression DNA construct.

38. A method of evaluating root architecture in a plant, comprising:

(a) introducing into a regenerable plant cell a suppression DNA construct comprising at least one regulatory element operably linked to:

(i) all or part of:

(A) a nucleic acid sequence encoding a polypeptide having an amino acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13, 17, 19, or 21, or

(B) a full complement of the nucleic acid sequence of (a)(i)(A); or

(ii) a region derived from all or part of a sense strand or antisense strand of a target gene of interest, said region having a nucleic acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to said all or part of a sense strand or antisense strand from which said region is derived, and wherein said target gene of interest encodes a RT1 or RT1-like polypeptide;

(b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the suppression DNA construct; and

(c) evaluating root architecture of the transgenic plant compared to a control plant not comprising the suppression DNA construct.

39. The method of claim 38, further comprising:

(d) obtaining a progeny plant derived from the transgenic plant, wherein the progeny plant comprises in its genome the suppression DNA construct; and

(e) evaluating root architecture of the progeny plant compared to a control plant not comprising the suppression DNA construct.

40. A method of evaluating root architecture in a plant, comprising:

(a) introducing into a regenerable plant cell a suppression DNA construct comprising at least one regulatory element operably linked to:

(i) all or part of:

(A) a nucleic acid sequence encoding a polypeptide having an amino acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 13, 17, 19, or 21, or

(B) a full complement of the nucleic acid sequence of (a)(i)(A); or

(ii) a region derived from all or part of a sense strand or antisense strand of a target gene of interest, said region having a nucleic acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to said all or part of a sense strand or antisense strand from which said region is derived, and wherein said target gene of interest encodes a RT1 or RT1-like polypeptide;

(b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the suppression DNA construct;

(c) obtaining a progeny plant derived from the transgenic plant, wherein the progeny plant comprises in its genome the suppression DNA construct; and

(d) evaluating root architecture of the progeny plant compared to a control plant not comprising the suppression DNA construct.

41. An isolated polynucleotide comprising:

(i) a nucleic acid sequence encoding a polypeptide having an amino acid sequence of at least 80%, sequence iden-

tity, based on the Clustal V method of alignment, when compared to SEQ ID NO:13 or 21; or

(ii) a full complement of the nucleic acid sequence of (i).

42. An isolated polynucleotide comprising:

(i) a nucleic acid sequence encoding a polypeptide having an amino acid sequence of at least 85%, sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO:13 or 21 or

(ii) a full complement of the nucleic acid sequence of (i).

43. An isolated polynucleotide comprising:

(i) a nucleic acid sequence encoding a polypeptide having an amino acid sequence of at least 90%, sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO:13 or 21; or

(ii) a full complement of the nucleic acid sequence of (i).

44. An isolated polynucleotide comprising:

(i) a nucleic acid sequence encoding a polypeptide having an amino acid sequence of at least 95%, sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO:13 or 21; or

(ii) a full complement of the nucleic acid sequence of (i).

45. The polynucleotide of claim 41, wherein the polypeptide sequence comprises SEQ ID NO:13 or 21.

46. The polynucleotide of claim 41, wherein the nucleic acid sequence comprises SEQ ID NO:12 or 20.

47. A vector comprising the polynucleotide of claim 41.

48. A recombinant DNA construct comprising the polynucleotide of claim 41 operably linked to at least one regulatory sequence.

49. A method for transforming a cell, comprising transforming a cell with the polynucleotide of claim 41.

50. A cell comprising the recombinant DNA construct of claim 48.

51. A method for producing a plant comprising transforming a plant cell with the polynucleotide of claim 41 and regenerating a plant from the transformed plant cell.

52. The isolated polynucleotide of claim 41, wherein the polypeptide sequence comprises at least one motif selected from the group consisting of SEQ ID NOs:22 and 23, wherein said motif is a substantially conserved subsequence.

53. An isolated polynucleotide encoding a polypeptide, wherein expression of said polypeptide results in an altered root architecture and wherein the polypeptide sequence comprises at least one motif selected from the group consisting of SEQ ID NOs:22 and 23, wherein said motif is a substantially conserved subsequence.

* * * * *